International Journal of Production Research

ISSN: 0020-7543 (Print) 1366-588X (Online) Journal homepage: https://www.tandfonline.com/loi/tprs20

Manufacturing data analytics using a virtual factory representation

Sanjay Jain, Guodong Shao & Seung-Jun Shin

To cite this article: Sanjay Jain, Guodong Shao & Seung-Jun Shin (2017) Manufacturing data analytics using a virtual factory representation, International Journal of Production Research, 55:18, 5450-5464, DOI: 10.1080/00207543.2017.1321799

To link to this article: https://doi.org/10.1080/00207543.2017.1321799

International Journal of Production Research, 2017
Vol. 55, No. 18, 5450–5464, https://doi.org/10.1080/00207543.2017.1321799

Manufacturing data analytics using a virtual factory representation

Sanjay Jain[a]*, Guodong Shao[b] and Seung-Jun Shin[b,c]

[a]*Department of Decision Sciences, School of Business, The George Washington University, Washington, DC, USA;*
[b]*Systems Integration Division, Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA;*
[c]*Graduate School of Management of Technology, Pukyong National University, Busan, South Korea*

(*Received 19 January 2016; accepted 11 April 2017*)

Large manufacturers have been using simulation to support decision-making for design and production. However, with the advancement of technologies and the emergence of big data, simulation can be utilised to perform and support data analytics for associated performance gains. This requires not only significant model development expertise, but also huge data collection and analysis efforts. This paper presents an approach within the frameworks of Design Science Research Methodology and prototyping to address the challenge of increasing the use of modelling, simulation and data analytics in manufacturing via reduction of the development effort. The use of manufacturing simulation models is presented as data analytics applications themselves and for supporting other data analytics applications by serving as data generators and as a tool for validation. The virtual factory concept is presented as the vehicle for manufacturing modelling and simulation. Virtual factory goes beyond traditional simulation models of factories to include multi-resolution modelling capabilities and thus allowing analysis at varying levels of detail. A path is proposed for implementation of the virtual factory concept that builds on developments in technologies and standards. A virtual machine prototype is provided as a demonstration of the use of a virtual representation for manufacturing data analytics.

Keywords: simulation applications; performance analysis; process modelling; CNC machining; production modelling; virtual factory; data analytics

1. Introduction

There have been multiple calls for increasing the use of Modelling and Simulation (M&S) for advancements in manufacturing. Use of M&S has been identified as one of the major steps in achieving smart manufacturing (SMLC 2012). A report on accelerating advanced manufacturing in United States calls for high fidelity M&S for reducing the design to manufacturing lead time and for advanced control and optimisation (PCAST 2014). Both these reports also call for increased use of data analytics for advancement in manufacturing. PCAST (2014) also notes links between analytics and simulation.

Data Analytics (DA) has been identified as a key to greater agility to react quickly to fluctuations in market demand or supplies as well as production control (Dean 2013). DA applications for manufacturing, referred to as Manufacturing Data Analytics (MDA), can be used to help improve manufacturing performance via insights into trends, patterns, areas of inefficiency and potential risks to manufacturers. Data for such analytics is increasingly available from advancing technology including sensors on machines and equipment, readers for radio-frequency identification tags and bar-codes, and data harvesting applications tracking information across multiple sources including financial transactions, market behaviour, and internet.

Currently, most of the manufacturing companies do not make good use of all the generated and collected data such as data from computer-aided design, computer-aided manufacturing and digital manufacturing production systems (Dean 2013). Those who are making an effort mainly analyse the data for improving and fine tuning elements such as business processes and customer service. MDA solutions can help collect, analyse and report the data dynamically. Many supply chain professionals including manufacturers are overwhelmed by the vast amount of data (Hazen et al. 2014). Most companies are struggling to get worthwhile returns from data analytics (Marchand and Peppard 2013). While many of the manufacturers are new to MDA, large manufacturers generally appear to have experience with M&S based on authors' interactions with industry personnel via multiple consortiums and conferences in United States and Europe. There is an opportunity to build on M&S use for MDA. This paper discusses multiple ways to use M&S to perform

*Corresponding author. Email: jain@gwu.edu

MDA and integration of M&S with other MDA applications for valuable decision support. In particular, the virtual factory concept is utilised as a representation of the real factory created using M&S for MDA.

The concept of virtual factory as a multi-resolution representation across the hierarchical levels of a real factory was proposed a number of years ago by Jain (1995) though other connotations of the phrase existed even earlier. Use of multi-resolution modelling can allow the flexibility to represent the components at varying levels of detail appropriate to the analysis of interest. Jain et al. (2001) attempted to create a multi-resolution virtual factory capability but found it highly challenging due to limitations of technology and information availability at the time. This paper identifies the technical advancements since then that make the implementation of the virtual factory vision within reach.

The primary contribution of the paper is the proposed path to implementation of the virtual factory concept and its use for MDA. The path and the implementation are demonstrated using a prototype virtual machine. In addition, relevant technologies and standards are identified to support the implementation of the virtual factory concept.

This work follows Design Science Research Methodology (DSRM) from the field of information systems (Peffers et al. 2008) and prototyping from operations management (Meredith et al. 1989). Simulation models for decision support have to be deployed as part of manufacturing information systems. Use of DSRM allows consistency with extant literature, guidance via a process model and template for presenting and evaluating design science in information systems. DSRM has six steps: problem identification, objectives definition, design and development, demonstration, evaluation and communication. Prototyping involves building an exemplar of a selected subset of the attributes of the system. Its use is motivated by the need to evaluate the proposed concept before embarking on a huge effort.

Following the discussion of the relevant efforts reported in the literature in the next section, Sections 3–8 are organised following the six DSRM steps discussed above. Sections 3 and 4 define the problem and objectives, respectively. Section 5, first, discusses the concepts and technologies employed for the design and development. This is followed by conceptual design of the virtual factory. The remaining steps of DSRM are carried out with a prototype of a virtual machine that is a unit component of the envisaged virtual factory and simulates the operations of a real machine. The prototype is evaluated for its applicability for use of M&S and MDA. The prototype presents a first step to implement the virtual factory concept. Section 9 concludes the paper with discussion of future work.

2. Related work

Related efforts are briefly reviewed in two areas, simulation-based MDA and virtual factory.

2.1 *Simulation-based manufacturing data analytics*

LaValle et al. (2011) report based on a survey of nearly 3000 industry personnel that senior executives are looking to make data-driven decisions based on scenarios and simulations. However, there are only a few examples reported of use of simulation-based MDA. Pegden (2011) contends that the role of simulation is expanding to exploit information and predict the impact of change at all levels of the business. Dudas, Hedenstierna, and Ng (2011) utilise stochastic simulation to generate data that is mined to identify relationships among decision variables and objectives to support decision-making. Zhou et al. (2012) used simulation models of manufacturing system together with search procedure to select and evaluate green production strategies. Ng et al. (2011) introduce a novel methodology that integrates the concept of innovisation with discrete-event simulation and data mining techniques for the analysis and optimisation of production systems. Applying data mining to the data sets generated from simulation-based multi-objective optimisation allows automatic or semi-automatic discovery and interpretation of the hidden relationships and patterns for optimal production systems design/reconfiguration. Melouk et al. (2013) employ simulation and a search space based optimisation approach for identifying improvements in inventory practices and process capabilities for steel manufacturing.

The few examples from the literature support the promise offered by simulation-based MDA, yet raises the question of reason behind limited application. It is highly likely that the effort involved in building detailed simulation models of manufacturing systems may be acting as a roadblock. Manufacturers need support for rapid development of detailed representations of their manufacturing systems using simulation. Such representations can be provided by a generic virtual factory that can be configured based on input data as discussed in this paper.

2.2 *Virtual factory*

The term virtual factory has been defined in multiple ways in the manufacturing research and application domains including as a high fidelity simulation, a virtual organisation, a virtual reality representation and an emulation facility (Jain and Shao 2014). This paper utilises the virtual factory definition as a high fidelity simulation of a manufacturing factory.

The virtual factory concept has been implemented by leading manufacturing companies. Ford Motor Company improves assembly line performance in its European facilities by evaluating and optimising the designs using virtual factory systems (IMT Staff 2013). Volvo Group Global (2014) validates changes using virtual factories before their implementation into actual plants and has the goal that 'by 2020 all major Volvo Group plants will be virtually tested before any major changes are done in the real world'. Major commercial software vendors support development of virtual factories via integrated solutions for product, process and system design, simulation, and visualisation (Tolio et al. 2013).

Virtual data management, automatic model generation, static and dynamic simulation, and integration and communication are paramount to realising a virtual factory (Choi, Kim, and Noh 2014). However, most software tools are, in general, not supplied with these capabilities making it a challenge to develop a virtual factory. There are efforts addressing different aspects of the challenge. To enhance conventional simulations for a virtual factory, Bal and Hashemipour (2009) use Product-Resource-Order-Staff Architecture for modelling controls while the Quest simulation tool models the physical elements. To integrate models and enhance communication, Hints et al. (2011) developed a software tool named Design Synthesis Module. Debevec, Simic, and Herakovic (2014) use a virtual factory model to test and improve the schedules before implementation in the real factories of small and medium size enterprises. For production planning, Terkaj, Tolio, and Urgo (2015) present an ontology for a virtual factory to aid planning decisions.

The recent concept of 'Industry 4.0' or the fourth industrial revolution from Germany includes Cyber-Physical Systems (CPS) as a key component. The function of CPS has been identified as monitoring physical processes and creating a virtual copy of the physical world to support decentralised decision-making (Mario, Tobias, and Boris 2015). The 'virtual copy' and 'virtual plant models' discussed in the context of Industry 4.0 closely match the concept of virtual factory discussed in this paper. The virtual factory can support the factory design stage and once the factory is built it can transition into a component of the CPS and support MDA and decision-making.

The need identified in the literature for simulation-based MDA can be ably met using the virtual factory representations as suggested by reported applications. Current applications generally utilise custom developments of virtual factories that require a large effort and expertise. The development of a capability to largely auto-generate virtual factories rapidly using data from real factories in standard formats as proposed in this paper will significantly reduce the effort and expertise requirement.

3. Problem identification

There is a recognition among leading initiatives for advancement of manufacturing that there are barriers to increased use of data analytics and simulation (SMLC 2012; PCAST 2014). Large data collection efforts and significant model development expertise are required for use of simulation and other data analytics applications. SMLC (2012) calls for 'lower cost barriers for applying advanced data analysis, modelling, and simulation in core manufacturing processes'. The problem of high cost and expertise barriers to use of simulation and data analytics in manufacturing is the identified target for our effort.

4. Objectives definition

The objective of our effort is to support manufacturers in employing M&S and MDA to improve their performance. The sub-objectives are:

- to develop a capability that allows largely auto-generation of a virtual factory model, and,
- to demonstrate the use of the model as a data analytics application itself and for supporting other data analytics application.

The presented approach is anticipated to help in the long term to increase the use of M&S and MDA in multiple ways. First, use of standard format interfaces will enable largely auto-generation of the virtual factory model thus lowering the cost barrier of using M&S. Johansson et al. (2007) have used the Core Manufacturing Simulation Data (CMSD) standard for generating majority of the model at the factory level. This approach will be extended to multiple levels of the virtual factory. Second, with the increased interest in MDA, the past familiarity with M&S, and the presentation of virtual factory as a MDA tool and a platform for other MDA tools, manufacturers may be motivated to use the virtual factory once available as an initial entry path to use MDA. Third, the multi-resolution modelling capability is expected to lead to better understanding and decisions and further motivate the use of M&S and MDA.

5. Concept design and development

This section discusses the factors supporting the proposed development to increase M&S and MDA applications for manufacturing followed by the presentation of the conceptual design of the virtual factory. The first subsection establishes the link between M&S and MDA. The second subsection identifies the recent advancements in M&S technology and interface standards. The third subsection presents the virtual factory concept that builds on advancements in technology and standards and is aimed at enabling rapid use of M&S and MDA.

5.1 *Simulation roles for MDA*

Manufacturers have used M&S to analyse design and operations for a long time. Simulation application includes analysis of output data to generate insights and hence simulation itself is a tool for DA. A number of statistical analysis tools are available for analysis of simulation input data and such tools are also DA applications. Thus, simulation is supported by DA applications. Simulation can be used to generate realistic data to support the evaluation of DA applications and for filling in missing data for use by DA. Simulation thus supports DA applications. Simulation roles for MDA hence can be grouped in two categories, as a MDA application and as a support application for other MDA applications. These roles are discussed in the following subsections.

5.1.1 *Simulation as a MDA application*

The Gartner Analytic Ascendancy Model (Laney and Kart 2012) defines four major applications of DA as shown in Figure 1. These include descriptive (what happened?), diagnostic (why did it happen?), predictive (what is likely to happen?) and prescriptive (how can we make it happen?) analyses. A number of techniques may be used for the four major applications including data mining, regression, Bayesian network analysis, classification and clustering algorithms, machine learning algorithms, optimisation and data envelopment analysis. Simulation is useful for three of the areas, namely, diagnostic, predictive and prescriptive as discussed below.

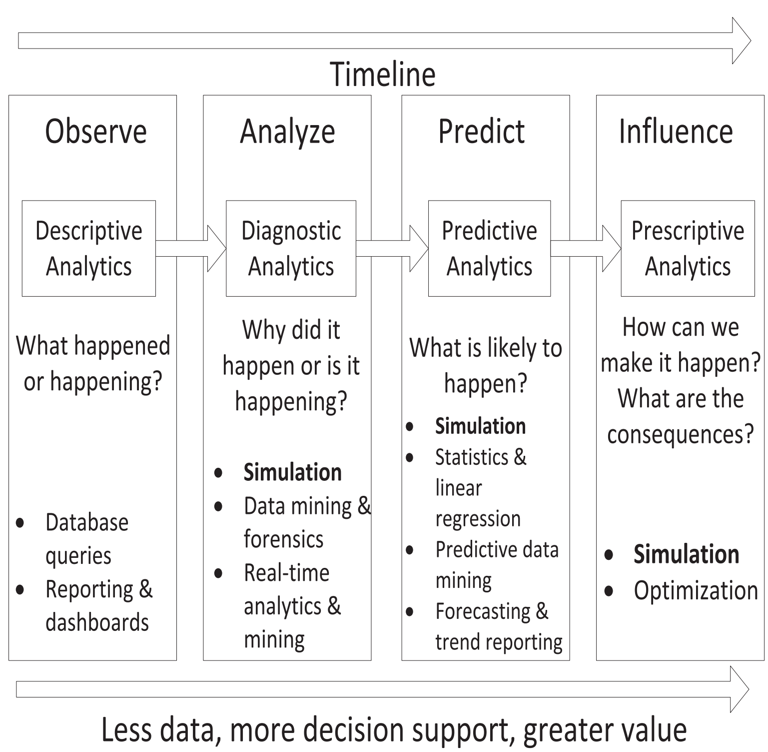

Figure 1. Role of simulation in major applications of data analytics.

5.1.1.1 *Simulation as a diagnostic analytics application.* Sensitivity analysis of simulation models allows identifying the factors that have an influence on the measure or patterns of interest. Such analysis can be used for diagnostic analytics. For example, one could investigate the causes of long cycle times for a certain product group by varying factors identified to be the potential contributors using a designed experiment. The results of such analysis can help answer the question of why the product group is experiencing long cycle times.

5.1.1.2 *Simulation as a predictive analytics application.* Predictive analytics is needed to understand the impact on performance measures of interest of future planned and unplanned changes such as in policies, product mix, resource availabilities and demands. Simulation can be used to effectively evaluate such impact via modelling of scenarios of planned and unplanned changes in a manufacturing system. It can thus answer the question of what is likely to happen with associated uncertainties.

5.1.1.3 *Simulation as a prescriptive analytics application.* Prescriptive analytics goes further than the predictive analytics and identifies the needed parameter settings and policies that will result in desired performance improvements. Simulation can contribute in three ways to the role of a prescriptive analytics application. First, simulation models can be exercised through a number of scenarios set up under a designed experiment to identify the parameter settings that improve performance measurements of interest. For example, parameters such as lot release rates, dispatching rules and resources per operation may be varied to identify the set that leads to improved due date performance.

Second, simulation can be used to take the output of another prescriptive analytics application such as an optimisation tool and fine tune it for implementation. It is generally highly cumbersome and at times not possible to include all the real-life constraints and variabilities in an optimisation model. For example, optimisation models can be used to represent the major constraints such as number of major resources (machines, operators) and technical precedences to generate a manufacturing schedule. Simulation can then be used to fine tune the solution using remaining real life constraints such as buffer spaces and transporters.

Third, simulation models can be interfaced with an optimisation procedure such as simulated annealing or genetic algorithm to iteratively search for the set of parameter settings that provide the best achievable performance. The third approach is generally referred to as a combined simulation optimisation approach. Also, combining simulation with optimisation as defined in the second and third approaches generally provides more accurate solutions than a standalone optimisation application.

5.1.2 *Simulation as a support application for other data analytics application*

Simulation outputs can be used to support other DA applications in multiple ways discussed below.

5.1.2.1 *Simulation as a data generator.* Simulation can be used to generate the data streams that can be used to test new DA applications. Manufacturing companies are generally understandably reluctant to provide access to data from their factories to researchers and developers of MDA applications. The proposed virtual factory models can be set up with hypothetical but realistic factory configurations to generate data streams that can be analysed by the DA applications under development. The virtual factory models should be able to generate data at the desired level of detail including machine level data streams for testing machine health DA applications and factory level data streams for testing factory performance DA applications.

5.1.2.2 *Simulation to support evaluation and validation.* It could be a challenge to ensure that the relationships identified by a MDA application are correct if the underlying true relationships are not known. Similarly, it could be a challenge to compare MDA applications based on the quality of their outputs if the corresponding true outputs are not known. Simulation can provide valuable support to meet both these challenges. Simulation scenarios can be prepared using known distributions for measures or patterns of interest and DA applications can be used to analyse outputs of simulation models to identify the particular patterns. For example, machine breakdown patterns analysed by a DA application being evaluated could be compared against the known distributions that were used in the simulation. In such use, simulation has an advantage over real manufacturing data since the underlying distributions may not be known accurately for the real data. Similarly, a manufacturer can use virtual factory representations of their specific manufacturing facilities with known distributions for aspects of interest to evaluate DA applications across multiple scenarios as part of an application selection exercise.

This subsection presented the multiple ways simulation can be used for MDA. In all the potential applications discussed, simulation models of the factory or its subsets are utilised as MDA applications themselves or in support of other MDA applications. A virtual factory representation of the real factory can hence serve to execute and support such application.

5.2 *Relevant technologies and standards*

Technologies relevant to the proposed virtual factory concept have been rapidly developing in recent years. This subsection briefly discusses technologies and standards that have now made the virtual factory concept implementation feasible.

5.2.1 *Simulation software technology*

The implementation of virtual factory concept requires contributions of component models from multiple research teams. Simulation software need to allow easy integration of such component models. Use of object-oriented programming can facilitate the integration of independently developed component models. About 15 to 20 years ago, there were few simulation software that offered the capability of object-oriented programming. Simple++ was one of the first object-oriented simulation software that was used for implementing a prototype of the virtual factory concept (Jain et al. 2001). At the time, all levels of resolution would have to be modelled in discrete event paradigm with Simple++ and all input data interfaces would have to be custom developed. It would have required a large effort to develop the virtual factory under those circumstances and hence it did not receive much interest.

Simulation software technology has developed at an accelerated rate similar to development in all fields of technologies. A few commercial simulation software are available currently that offer object-oriented development environments and allow mixing multiple simulation paradigms in the same model. These software facilitate development and integration of multi-resolution models. An example is the integration of a system dynamics model of a supply chain with a discrete simulation model of one of the manufacturing nodes using a commercial simulation software (Jain et al. 2013). It is possible to have component models developed by different teams using the same software integrated together in one executable file. However, integration of models developed in different simulation software still needs to utilise distributed simulation.

5.2.2 *Distributed simulation technology*

As simulation environments for manufacturing evolve to be more collaborative, open, and global, distributed simulation technology is anticipated to become essential for implementing virtual factory concepts. The first standard mechanism for distribution simulation was the High Level Architecture (HLA) (Kuhl, Weatherly, and Dahmann 1999) that required advanced expertise and large effort for implementation. More recently, researchers are using web services technology for integration of distributed simulation (Yoo et al. 2009). Recent updates in the HLA standard include web services support (IEEE 2010). The recently developed Standard for Commercial Off-the-Shelf Simulation Package Interoperability (SISO 2010) helps identify problems and evaluate approaches for integration of models developed using different software. These recent developments have significantly facilitated the development of distributed simulation arrangements. A web services based approach is envisaged for implementation of distributed simulation for realising the virtual factory concept on a wider scale.

5.2.3 *Data interface standards*

The implementation of virtual factory will be facilitated by standard data interfaces that allow reading in and generating data using the same formats as those used in a real factory. A wide range of formats are used for the data streams generated by a real factory in practice and that poses a challenge. Efforts have been made in recent years to develop standards for some of these interfaces including those identified below. Standards like these are being considered for implementation of the open-standards-based interfaces identified in Figure 2 that is discussed in the next section.

- Open Applications Group Integration Specification (OAGIS) is a cross industry canonical model for defining business messages. It is open standards based and uses eXtensible Markup Language (XML). It also addresses application-to-application and business-to-business integration via business processes (Connelly and Hertlein 2010).
- Core Manufacturing Simulation Data (CMSD) defines neutral interface for sharing data between manufacturing applications and simulation (SISO 2012). A number of case studies have used the CMSD model (e.g., Johansson et al. [2007]).

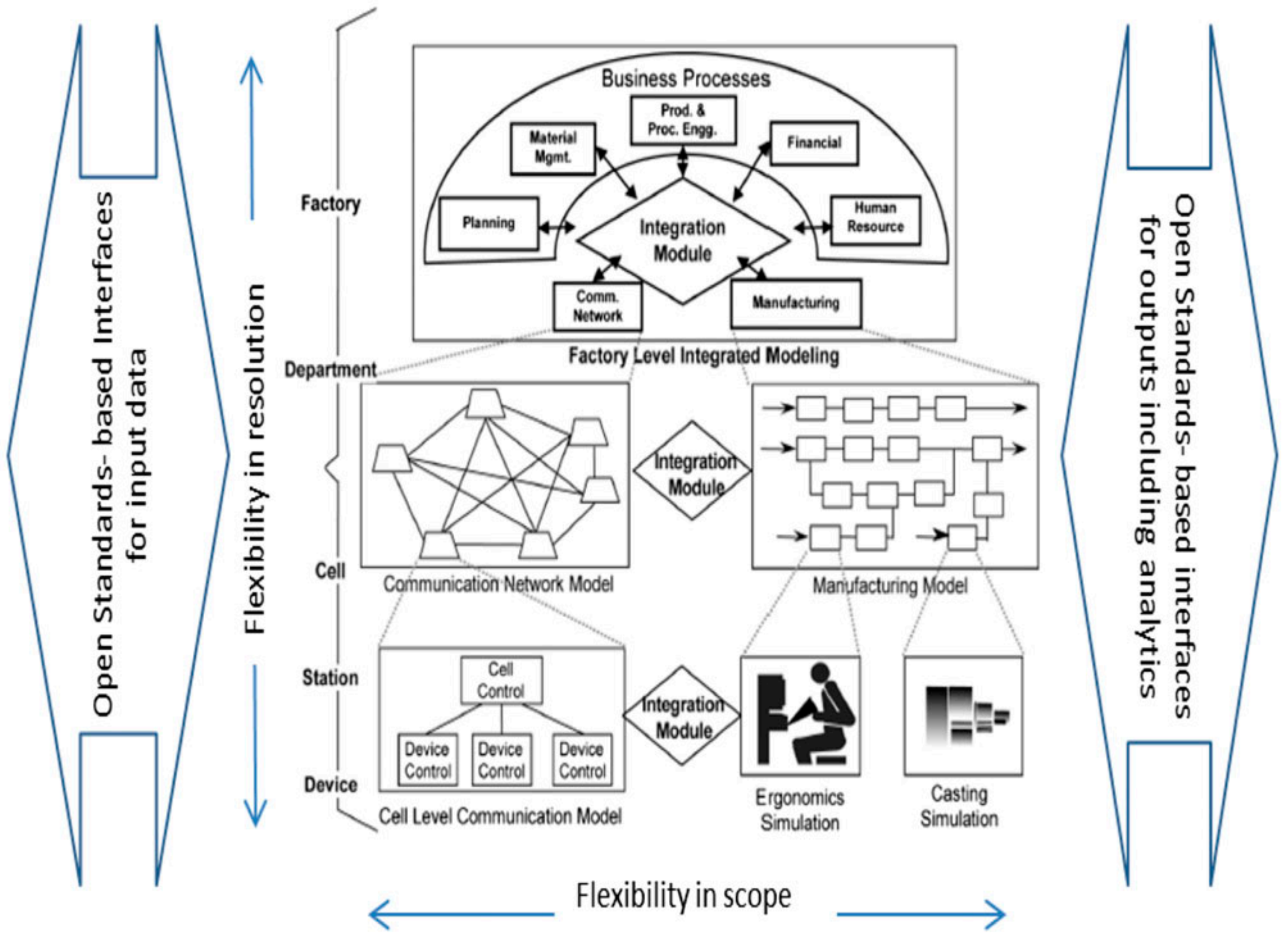

Figure 2. Virtual factory concept (adapted from Jain et al. [2001] and enhanced).

- ISA-95 developed by International Society for Automation (ISA) provides models and terminology to identify the information to be exchanged between manufacturing systems (ANSI 2010).
- Business To Manufacturing Markup Language (B2MML) implements the ISA-95 data models through a set of XML schemas. B2MML allows businesses to easily integrate their Enterprise Resource Planning (ERP) system modules with their Manufacturing Execution System solutions.
- STEP-compliant data interface for Numerical Controls (STEP-NC) is a standard for representing process planning data. STEP-NC uses the concept of working step to specify machining processes (IMS STEP-NC Consortium 2003).
- MTConnect ('MT' appears to refer to Machine Tools) is a XML based standard for extracting data from numerically controlled machine tools (AMT 2014).

In the above list, CMSD specifically includes information to support simulation models of manufacturing while the rest of them are designed for operational use and can be exploited for realising virtual factories. The use of two of the standards in the list, STEP-NC and MTConnect, for the virtual machine prototype in Section 6 demonstrates how they can be used in implementation of the virtual factory concept.

5.3 *Virtual factory concept*

The virtual factory concept is presented in the context of its use as and in support of MDA in this subsection. Indeed the recent high interest in MDA prompted another look and enhancement of the virtual factory concept defined a while ago by Jain et al. (2001). Virtual factory can support MDA beyond the real factory since it allows generating hypothetical normal and extreme scenarios for deeper understanding of consequences of different strategies and decisions.

The current focus of the effort is on modelling discrete manufacturing operations ranging from job shops to flow shop configuration. The virtual factory should model the primary production process and associated supporting processes including materials management, demand management, product and manufacturing engineering, and maintenance. It should have multi-resolution modelling capability, that is, the ability to model all the processes at multiple levels of detail with the level to be selected collectively or individually based on the problem to be analysed. The virtual factory model with multi-resolution capability mimics the real factory as it allows analysis at the level of detail selected by analysts and decision-makers.

The multi-resolution modelling capability should be built on the ability to employ modelling paradigms appropriate to the resolution level. The flow of parts through multiple production stages is generally best modelled using discrete event simulation. Agent-based simulation may be recommended way to model the behaviour of individual machines and operators. Modelling of machine component movements and associated forces may best be done using continuous simulation of physical-sciences-based equations. Users should be able to select the paradigm for different component models, validate the component models individually, integrate them together and validate the integrated models.

The multi-resolution capability may be utilised in two ways. First, all parts of the model may be set at a homogenous level of detail. For example, the material flow may be modelled across all parts of the factory at low resolution level, or it may be modelled at a high resolution level though that may be rarely found useful. Second, parts of the model may be at heterogeneous levels of detail with the part of primary interest at a higher level of detail than others. Analysts have been generally used to models with a single level of detail in the past and hence should be able to seamlessly move to use of the virtual factory model with a homogenous level of detail. An example of such use is employing a discrete event simulation model to analyse the cycle times of various product types based on defined resource configuration and operating policies.

Using the virtual factory model with its parts at heterogeneous levels of detail will take some additional effort to set up, but it will provide for a more efficient use of analyst and computer time. The model set up should be based on swapping of model components with one level of detail with a corresponding one with a higher level of detail. As an example, a model with heterogeneous detail levels may be used to analyse an issue with the performance of a tightly connected sub-assembly line in a factory. The virtual factory model may be set up to represent the sub-assembly line at a higher level of detail than the rest of the factory. The sub-assembly line model should include representation of moves within each work station to allow identifying issues within work stations contributing to the line's performance. The rest of the factory may be modelled without the detail inside the work stations with the purpose of modelling common resources such as material handling, modelling flow of needed materials to the sub-assembly line, and for modelling flows absorbing the output of the sub-assembly line at realistic rates. Modelling of rest of the factory in an integrated manner with the detailed sub-assembly line model allows higher accuracy compared to the alternative of using only the sub-assembly line model with distributions for arrival of materials and removal of products and assumptions on common resource availabilities. It offers improved efficiency for analyst and computer execution compared to the alternative of modelling the entire factory at high level of detail as it saves the time required to verify and import detailed data for entire factory and for executing a much larger model.

A generic capability should be developed to allow the virtual factory model being largely auto-generated using detailed data in standard formats from a real or a designed factory. The virtual factory model for a factory will be essentially generated by instantiating the model with factory specific data. This will thus reduce the effort and the associated cost required to implement virtual factories at successive factories and allow a large number of manufacturers to benefit. Input files with data on factory and products configuration should be based on available standards as much as possible. The CMSD standard included in the list in the preceding subsection is a good candidate for such configuration data at factory flow level of detail that is typically modelled using discrete event simulation.

The summary and streaming outputs of the virtual factory model should also comply with available standards as far as possible. The user should be able to select a standard if multiple standards can be used for any of the output information. Among the standards listed in the preceding subsection, B2MML is a good candidate for reporting manufacturing performance data at factory level while MTConnect is a good candidate for generating machine level data streams. In addition to standard formats, capabilities should be available to customise the output formats.

Previously the virtual factory concept focused on the multi-resolution representation of a real factory. Availability and use of open-standards based interfaces can provide a critical push to development and implementation of the virtual factory concept. The ability to easily vary the scope and resolution levels of component models should provide a considerable benefit to MDA applications. The enhanced concept of virtual factory is shown in Figure 2. The resolution levels shown in Figure 2 traverse the manufacturing system hierarchy using factory specific terms and are in close agreement with descriptions of such hierarchy in IEC 62264-3 standard (IEC 2007).

Admittedly, it will take a joint effort of multiple teams of highly skilled simulation researchers to further develop and implement the virtual factory concept as described here. The motivation for taking on such a large task comes from the expectation that once the capability is available it can lead to a step jump in the use of simulation and DA in manufacturing. A good approach to develop the large capability is to proceed in a bottom-up manner starting with individual component models at the machine level.

6. Demonstration via a virtual machine prototype

As a first step towards the proposed virtual factory, a virtual machine prototype is developed to generate machine level data streams that can be used by a diagnostic analytics application. The conceptual design presented in the previous section is being implemented in a bottom up manner and has been initiated with virtual machines capable of accurately simulating operations of real machines. The verified and validated virtual machines can be integrated together as component models into a virtual factory model. Factory level streams can be generated using an aggregation of machine level streams generated by component virtual machines.

This section presents the virtual machine as a demonstration of a component of the virtual factory concept. The first subsection introduces the design and implementation of the virtual machine and the second subsection describes the extension of these virtual machines towards constructing a virtual work cell.

6.1 *Virtual machine*

The virtual machine has been developed using STEP-NC to MTConnect (STEP2M) simulator. The STEP2M simulator simulates machining processes using process planning data as input. It utilises physical-sciences-based equations to model the energy consumption during the machining process. It generates machine performance data that can be fed into MDA applications.

6.1.1 *Design of virtual machine*

The simulator is composed of three major modules (1) STEP-NC processing, (2) machining estimation and (3) MTConnect generation. The functional architecture of the simulator including simulation functions and data flows is shown in Figure 3.

The virtual machine simulator uses three main inputs. First, the capabilities of the machine are defined using the machine tool specification. Second, the code scheme for a G-code programme, i.e. Numerical Control (NC) programme, is described using the NC system. Third, the STEP-NC programme defines the process to be executed. The simulator streams machine performance data in MTConnect standard format as its output.

The real NC machines process a part by translating a STEP-NC programme into machine-interpretable format. This is modelled in the simulator using the STEP-NC processing module. First, the STEP-NC interpretation step parses the STEP-NC programme to instantiate the STEP-NC objects using the defined data scheme. Next, the STEP-NC objects are used by the tool path generation step to create a tool path with sequential tool movements and their instructions for rapid or interpolation trajectory. The final step of the module, G-code generation, produces a G-code programme that defines actions such as tool selection, spindle, and feed rates, and provides the tool path.

The G-code programme is used by the machining estimation module to determine movements and power consumption over time. The event/time/position estimation step models time-dominant events experienced by the machine components and uses this information to determine corresponding tool positions. The power estimation steps calculate the power required to accomplish the machine component actions. The estimated events, tool positions and consumed power over time are provided to the MTConnect generation module.

The MTConnect generation module creates the data streaming document based on MTConnect standard on demand. Three steps are involved in this process (Shao, Jain, and Shin 2014): (1) machine specification registration, (2) collection of runtime data and (3) MTConnect data request. Further details on the STEP2M simulator are available in Shin et al. (2016).

6.1.2 *Implementation of virtual machine*

A virtual machine prototype has been built for a two-axis turning machine using the architecture described in the previous sub-section. The prototype is developed in Java and utilises PrimeFaces for a web interface and Tomcat for a MTConnect server. The machine tool specification and G-code instructions are defined using a turning machine tool and a FANUC 0-series controller.

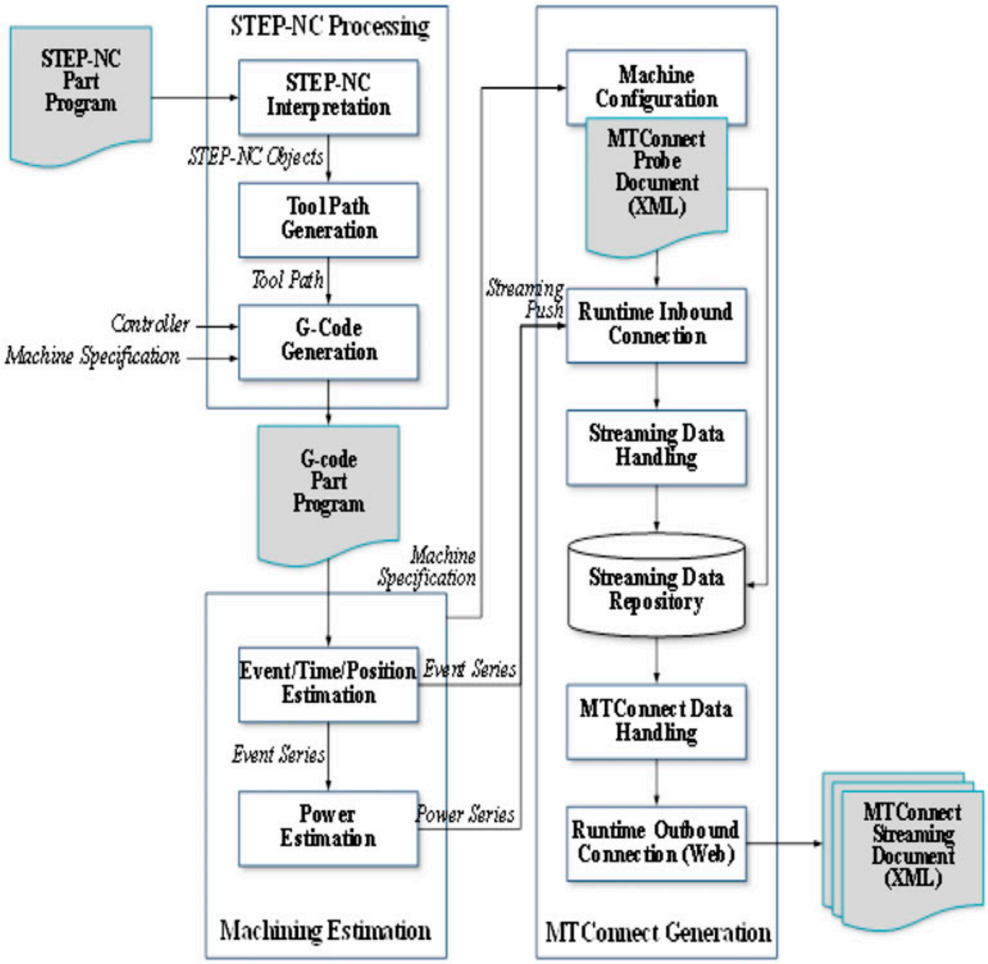

Figure 3. Functional architecture of STEP2M simulator.

The prototype generates an MTConnect XML file as an output. The MTConnect standard format includes fields for device, components and data items. The device is the machine tool generating the stream. The components are the main physical modules of the machine tool including axes (X, Z and rotary), the coolant system, and the main body. The data items include the monitoring data over time such as the tool position and the power consumed by the defined components.

The accuracy of the virtual machine is validated via comparison of simulated results with actual measurements. Time series data for simulated and actual power consumptions presented in Figure 4 shows considerable agreement. The variations in the time series data marked by numbers 1 through 4 in Figure 4 were analysed as follows: (1) a momentary power peak occurs due to inertia to make the spindle start rotating; (2) when the coolant system is turned on there is a step up in the power use; (3) the tool contact with the workpiece results in a small peak followed by the cutting power that reduces with successive cuts corresponding to decreasing workpiece diameter; and (4) there is a reverse peak calculated in simulated power by the machining estimation module for stopping the spindle while the actual machine stops with natural deceleration without requiring the reverse power.

6.2 *Towards a virtual work cell*

The purpose of implementing the STEP2M simulator above is to construct a virtual factory environment where such simulator can be used as a component model at a detailed resolution level. The component models need to be integrated successively into higher level models for this purpose. A virtual machine has been integrated into a virtual work cell using agent-based simulation capabilities of AnyLogic, a commercial simulation software (Lechevalier et al. 2015). Figure 5 shows the behaviour model of the virtual cell with a single virtual machine representing a milling machine.

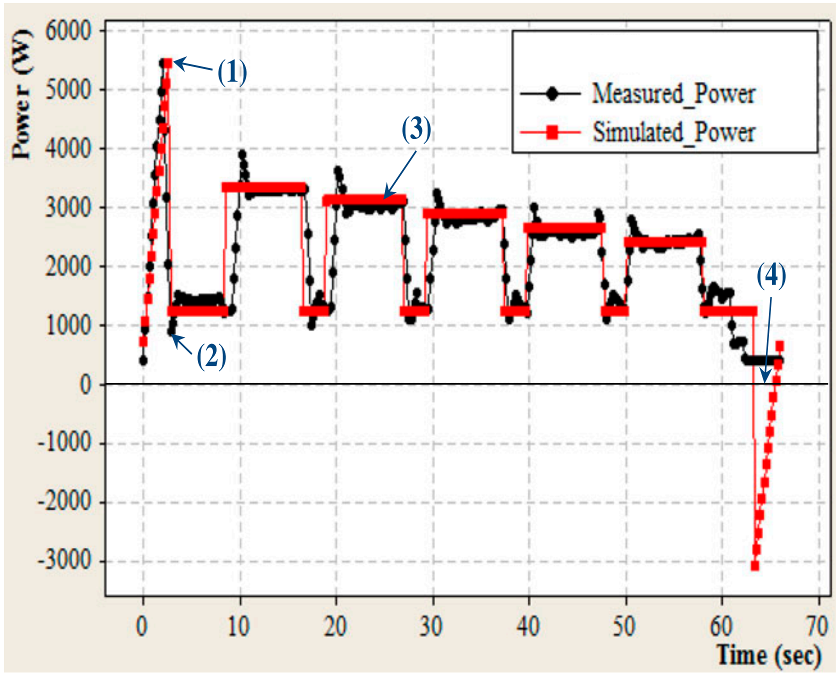

Figure 4. Comparison of measured and simulated power over time.

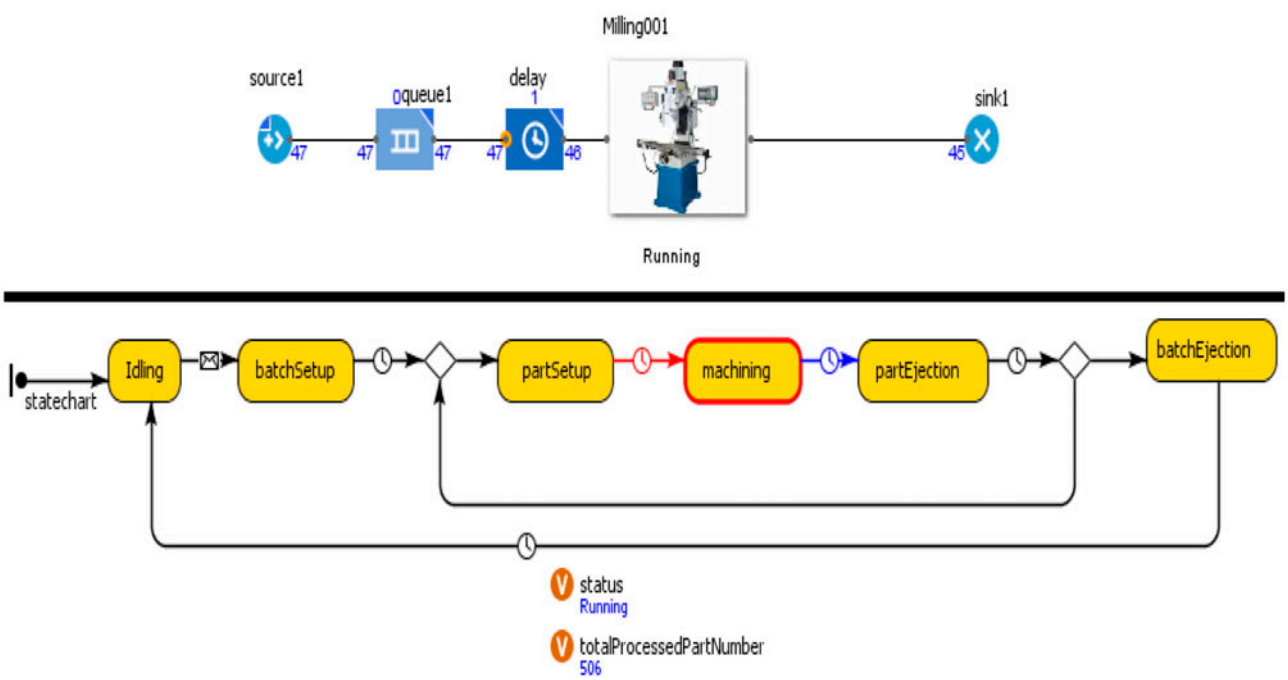

Figure 5. An example of a virtual milling machine's behaviour model (adapted from Lechevalier et al. [2015]).

The upper part of the figure represents the part flow in the cell while the lower part shows the machine behaviour using a state diagram. This model can generate several performance metrics for the work cell, including time spent in idling, batch set-up, part set-up, machining (modelled using STEP2M simulator), part ejection and batch ejection, as well as the machine-level data streams presented in Figure 4. When multiple virtual machines are modelled and installed for the designated work cell, this simulation capability can provide multi-resolution modelling specifically for the device/station and cell levels presented in Figure 2. The cell level is modelled at a lower resolution using discrete event simulation capabilities of AnyLogic and the virtual machines are modelled using agent-based simulation capabilities of AnyLogic for higher resolution modelling. Proceeding in such a bottom up manner, the holistic virtual factory environment proposed in Section 5 can thus become feasible.

7. Evaluation: role of the virtual machine in data analytics

The virtual machine prototype utilises M&S for MDA application as envisaged in the design and development section albeit at a much smaller scale than the envisioned virtual factory. The virtual machine prototype can support the simulation roles for MDA discussed earlier in Section 5.1 specifically as a diagnostics analytics application (Section 5.1.1.1) and as a data generator (Section 5.1.2.1).

Traditional diagnostics analyses for machine performance using simulation typically use a NC programme as input. This limits the analyses since the NC programme alone is not sufficient to determine the process sequence and parameters for process planning. This limitation is removed in the presented virtual machine prototype using the process plan data as input in place of a NC programme. Employing the STEP-NC standard allows use of object-oriented working steps for precise specification of the process sequence and selected parameters. The output machine performance data corresponds to the provided input. The virtual machine prototype can iteratively generate a set of input and output data that can be used for diagnostics analytics thus meeting the role of simulation discussed in Section 5.1.1.1.

Machining monitoring data including tools status, events and movements have to be captured to assess the efficiency of machining operations via data analytics (Muchiri and Pintelon 2008). Collecting such data on a real machine requires additional effort and costs due to the need for placing measurement sensors and associated interfaces. The use of the virtual machine prototype obviates the need for such physical measurement devices. It generates the monitoring data needed for analytics thus meeting the role of simulation as a data generator discussed in Section 5.1.2.1. The current prototype has been set up to generate power consumption and machining time measures. Other metrics, such as surface roughness and machining error, are planned for future.

There are some limitations in reproducing real phenomena perfectly as indicated in Figure 4. These limitations may arise from a lack of real data to be compared and limitations of simulation itself. These limitations may be overcome with efforts for developing more realistic models and calibrating data by adjusting and tuning simulated data with referential real data.

The virtual machine prototype demonstrates the feasibility of generating machine-level data streams. The provided example of integrating the virtual machine into a virtual cell demonstrates the feasibility of bottom up approach for building the virtual factory. The prototype also demonstrates the feasibility of bringing together M&S and MDA capabilities. It provides confidence for building a virtual factory by developing and integrating several virtual machines with the identification of different machine specifications and production planning data. Different machine specifications will allow modelling of individual performances of the machines. Production planning data will allow modelling of virtual machines' operations following a schedule and calculating factory-level metrics such as cycle times and resource utilisations.

8. Communication

This step of the DSRM focuses on communicating the problem and the solution together with its evaluation to researchers and practicing professionals. The problem has been recognised and has received attention in initiatives focusing on advancements in manufacturing. A few different solution efforts are in progress as reported in Section 2. Given the complexity of the problem, it behoves the researchers to mount a joint effort. Communication of the relevant developments will hopefully help bring the corresponding researchers together.

Publication of this paper in this leading scholarly journal is a key part of the execution of the communication step. Parts of the concept and its development have been also presented earlier at conferences that attract researchers and practicing professionals. The development and communication efforts are intended to continue towards realisation of the virtual factory model.

9. Conclusion

This paper supports the progress towards envisioned smart manufacturing systems via increased use of M&S and MDA to provide decision support. The effort is presented in the framework of DSRM from the information system field and prototyping from operations management field. Multiple ways are identified in which simulation can support MDA.

First, simulation can itself serve as an MDA application and used for diagnostic, predictive and prescriptive analyses. Second, simulation can support other MDA applications in two primary ways, (a) by generating the data comparable to that provided by real manufacturing systems, and (b) by providing the means to evaluate and validate the MDA applications. However, simulation and MDA require significant effort and expertise. Virtual factory modelling is proposed as the primary vehicle to apply M&S and MDA and reducing the barriers to their application. The effort for developing virtual factories can be reduced by utilising standard driven approaches to auto-generate the models using the data from real factories. A first step in the bottom up approach to develop the virtual factory is the creation of a prototype virtual machine. The evaluation of virtual machine indicates that it meets the envisaged roles of simulation as a diagnostic analytics application and as a data generator to support MDA applications.

This paper makes contributions in two areas, virtual factory and MDA. It presents an enhanced version of the virtual factory concept and associated interfaces in the context of MDA. In addition, the paper provides a path to implementation of a generic virtual factory building on the advancements in technologies for simulation software and distributed simulation and relevant interface standards. In the area of MDA, the paper presents multiple ways in which the virtual factory can serve as a MDA application or support other MDA applications. The initial results from the prototype evaluation are shared to motivate further development of the virtual factory capability by the research community.

Future research directions include development of additional machine models for different processes, development of models at higher level of the manufacturing system hierarchy, and integration of component models across the hierarchy. While the near term directions are focused on developing the virtual factory component models within one software environment for ease of integration, distributed simulation arrangements for integrating component models developed in different software are anticipated in longer term. Near term objectives also include developing and interfacing a prototype virtual cell model to an MDA application.

Disclaimer

No approval or endorsement of any commercial product by the National Institute of Standards and Technology (NIST) 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.

Acknowledgement

The authors would like to thank the editors and anonymous reviewers for their valuable and constructive suggestions that led to considerable improvement of this paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the United States Government; the cooperative agreement [grant number 70NANB13H158] between NIST and George Washington University; and NIST's Foreign Guest Researcher Program. This research was partially supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) [grant number 2016R1C1B1008820].

References

AMT. 2014. *MTConnect Standard Part 1 – Overview and Protocol Version 1.3.0*. AMT – The Association for Manufacturing Technology, September 30, 2014.

ANSI. 2010. *ANSI/ISA-95.00.01: Enterprise-control System Integration – Part 1: Models and Terminology*. Washington, DC: American National Standards Institute.

Bal, M., and M. Hashemipour. 2009. "Virtual Factory Approach for Implementation of Holonic Control Industrial Applications: A Case Study in Die-casting Industry." *Robotics and Computer-Integrated Manufacturing* 25: 570–581.

Choi, S., B. H. Kim, and S. D. Noh. 2014. "A Diagnosis and Evaluation Method for Strategic Planning and Systematic Design of a Virtual Factory in Smart Manufacturing Systems." *International Journal of Precision Engineering and Manufacturing* 16: 1107–1115.

Connelly, D., and R. Hertlein. 2010. "An Open Standards Based, Cross Industry Canonical Model." Open Applications Group. http://www.oagi.org/oagi/downloads/ResourceDownloads/2011_0428_OAGIS_Canonical.pdf.

Dean, B. 2013. "Data in Motion: The Next Frontier for Manufacturers." Manufacturing.net. http://www.manufacturing.net/articles/2013/08/data-in-motion-the-next-frontier-for-manufacturers.

Debevec, M., M. Simic, and N. Herakovic. 2014. "Virtual Factory as an Advanced Approach for Production Process Optimization." *International Journal of Simulation Modeling* 13 (1): 66–78.

Dudas, C., P. Hedenstierna, and A. H. Ng. 2011. "Simulation-based Innovization for Manufacturing Systems Analysis Using Data Mining and Visual Analytics." Proceedings of the 4th Swedish Production Symposium, 374–382.

Hazen, B. T., C. A. Boone, J. D. Ezell, and L. A. Jones-Farmer. 2014. "Data Quality for Data Science, Predictive Analytics, and Big Data in Supply Chain Management: An Introduction to the Problem and Suggestions for Research and Applications." *International Journal of Production Economics* 154: 72–80.

Hints, R., M. Vanca, W. Terkaj, E. D. Marra, S. Temperini, and D. Banabic 2011. "A Virtual Factory Tool to Enhance the Integrated Design of Production Systems." In Proceedings of the DET2011 7th International Conference on Digital Enterprise Technology, Athens, Greece, 28–30.

IEC. 2007. *IEC 62264-3 ed1.0 Enterprise-control System Integration – Part 3: Activity Models of Manufacturing Operations Management*. International Electrotechnical Commission, Switzerland.

IEEE. 2010. *1516.1-2010 IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) – Federate Interface Specification*. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

IMS STEP-NC Consortium. 2003. *STEP-compliant Data Interface for Numerical Controls (STEP-NC)*. Technical Report 3 of IMS Project (97006), IMS STEP-NC Consortium, Erlangen, Germany, January 1–June 31.

IMT Staff. 2013. "Are Virtual Factories the Future of Manufacturing?" *Thomasnet Industry News*, March 13. http://news.thomasnet.com/IMT/2013/03/13/are-virtual-factories-the-future-of-manufacturing/.

Jain, S. 1995. "Virtual Factory Framework: A Key Enabler for Agile Manufacturing." *Proceedings of 1995 INRIA/IEEE Symposium on Emerging Technologies and Factory Automation, Paris* Oct. 1: 247–258.

Jain, S., and G. Shao. 2014. "Virtual Factory Revisited for Manufacturing Data Analytics." *Proceedings of the 2014 Winter Simulation Conference*, 887–898. Piscataway, NJ: IEEE.

Jain, S., N. Choong, K. Aye, and M. Luo. 2001. "Virtual Factory: An Integrated Approach to Manufacturing Systems Modeling." *International Journal of Operations & Production Management* 21 (5/6): 594–608.

Jain, S., E. Lindskog, J. Andersson, and B. Johansson. 2013. "A Hierarchical Approach for Evaluating Energy Trade-offs in Supply Chains." *International Journal of Production Economics* 146 (2): 411–422.

Johansson, M., B. Johansson, A. Skoogh, S. Leong, F. Riddick, Y. T. Lee, G. Shao, and P. Klingstam. 2007. "A Test Implementation of the Core Manufacturing Simulation Data Specification." *Proceedings of the 2007 Winter Simulation Conference*, 1673–1681. Piscataway, NJ: IEEE.

Kuhl, F., R. Weatherly, and J. Dahmann. 1999. *Creating Computer Simulations: An Introduction to the High Level Architecture*. Upper Saddle River, NJ: Prentice Hall.

Laney, D., and L. Kart. 2012. *Emerging Role of the Data Scientist and the Art of Data Science*. Report G00227058. Stamford, CT: Gartner, Inc.

LaValle, S., E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz. 2011. "Big Data, Analytics and the Path from Insights to Value." *MIT Sloan Management Review* 52 (2): 21.

Lechevalier, D., S.-J. Shin, J. Woo, S. Rachuri, S. Foufou. 2015. "A Virtual Milling Machine Model to Generate Machine-monitoring Data for Predictive Analytics." *Proceedings of the 12th IFIP WG 5.1 International Conference (PLM 2015*), Doha, Qatar, October 19–21: 835–845.

Marchand, D. A., and J. Peppard. 2013. "Why IT Fumbles Analytics." *Harvard Business Review* 91 (1): 104–112.

Mario, H., P. Tobias, and O. Boris. 2015. "Design Principles for Industrie 4.0 Scenarios: A Literature Review." Working Paper No. 01 / 2015, Technische Universität Dortmund, Dortmund, Germany.

Melouk, S., N. Freeman, D. Miller, and M. Dunning. 2013. "Simulation Optimization-based Decision Support Tool for Steel Manufacturing." *International Journal of Production Economics* 141 (1): 269–276.

Meredith, J. R., A. Raturi, K. Amoako-Gyampah, and B. Kaplan. 1989. "Alternative Research Paradigms in Operations." *Journal of Operations Management* 8 (4): 297–326.

Muchiri, P., and L. Pintelon. 2008. "Performance Measurement Using Overall Equipment Effectiveness (OEE): Literature Review and Practical Application Discussion." *International Journal of Production Research* 46 (13): 3517–3535.

Ng, A. H. C., C. Dudas, J. Nießen, and K. Deb. 2011. "Simulation-based Innovization Using Data Mining for Production Systems Analysis." *Multi-objective Evolutionary Optimisation for Product Design and Manufacturing* 401–429. doi:10.1007/978-0-85729-652-8_15.

PCAST. 2014. Steering Committee Report-Accelerating US Advanced Manufacturing: Annex 1–10: Manufacturing Technologies. President's Council of Advisors on Science and Technology (PCAST). Accessed August 5, 2016. https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/amp2.0-annex1-10-manufacturingtechnologyareas.pdf

Peffers, K., T. Tuunanen, M. A. Rothenberger, and S. Chatterjee. 2007. "A Design Science Research Methodology for Information Systems Research." *Journal of Management Information Systems* 24 (3): 45–77.

Pegden, C. D. 2011. "Simulation-based Performance Analytics." Accessed 15 June 2016. http://www.simio.com/resources/white-papers/Simulation-based-performance-analysis/Simulation-based-performance-analytics.pdf

Shao, G., S. Jain, and S.-J. Shin. 2014. “Data Analytics Using Simulation for Smart Manufacturing.” *Proceedings of the 2014 Winter Simulation Conference*, 2192–2203. Piscataway, NJ: IEEE.

Shin, S.-J., J. Woo, D. Kim, S. Kumaraguru, and S. Rachuri. 2016. “Developing a Virtual Machining Model to Generate MTConnect Machine-monitoring Data from STEP-NC.” *International Journal of Production Research* 54 (15): 4487–4505.

SISO. 2010. *SISO-STD-006-2010: Standard for Commercial Off-the-Shelf (COTS) Simulation Package Interoperability (CSPI) Reference Models*. Orlando, FL: Simulation Interoperability Standards Organization.

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

SMLC 2012. SMLC Forum: Priorities, Infrastructure, and Collaboration for Implementation of Smart Manufacturing: Workshop Summary Report. Smart Manufacturing Leadership Coalition (SMLC). Washington, DC. Oct. 2–3. Accessed August 5, 2016. https://smartmanufacturingcoalition.org/sites/default/files/smlc_forum_report_vf_0.pdf

Terkaj, W., T. Tolio, and M. Urgo. 2015. “A Virtual Factory Approach for in Situ Simulation to Support Production and Maintenance Planning.” *CIRP Annals – Manufacturing Technology* 64: 451–454.

Tolio, T., M. Sacco, W. Terkaj, and M. Urgo. 2013. “Virtual Factory: an Integrated Framework for Manufacturing Systems Design and Analysis.” *Proceedings of Forty Sixth CIRP Conference on Manufacturing Systems*, 25–30. Amsterdam: Elsevier.

Volvo Group Global. 2014. “Step into the Virtual Factory.” Accessed January 19, 2016. http://www.volvogroup.com/group/global/en-gb/researchandtechnology/Sustainable_production/Virtual_manufacturing/Pages/virtual_manufacturing.aspx.

Yoo, T., K. Kim, S. Song, H. Cho, and E. Yücesan. 2009. “Applying Web Services Technology to Implement Distributed Simulation for Supply Chain Modeling and Analysis.” *Proceedings of the 2009 Winter Simulation Conference*, 863–873. Piscataway, NJ: IEEE.

Zhou, M., Y. Pan, Z. Chen, W. Yang, and B. Li. 2012. “Selection and Evaluation of Green Production Strategies: Analytic and Simulation Models.” *Journal of Cleaner Production* 26: 9–17.