# A DATA DRIVEN APPROACH TO AUTOMATED SIMULATION MODEL BUILDING

Charu Chandra
Industrial & Manufacturing Systems Engineering
University of Michigan-Dearborn
4901 Evergreen Road
Dearborn, MI 48128, USA
E-mail: charu@umich.edu

Jānis Grabis
Department of Operations Research
Riga Technical University
Kalku 1
Riga, LV-1658, Latvia
E-mail: grabis@itl.rtu.lv

## KEYWORDS

Automated model building, simulation modeling, multi-stage manufacturing system, data model.

## ABSTRACT

This paper develops a simulation model building approach aimed at reducing model development efforts. The model building approach is intended for modeling of multi-stage manufacturing systems. A data model is used for representing the manufacturing system in the standardized manner. The data model is created from multiple raw data sources. A simulation model is automatically generated on the basis of the predefined template using information provided in the data model. Different manufacturing systems can be modeled by changing information in the data model. The automated generation allows avoiding model building errors caused by the large scale of the modeling problem. The model building approach is applied to study scheduling at initial stages of the manufacturing process of an automotive company. The automated approach is suitable for the problem because the system contains large number of similar objects and the company operates several similar systems.

## INTRODUCTION

High degree of complexity is the characteristic of a majority of large manufacturing systems. The complexity is caused by interactions among multiple-products, production stages and processing technologies, and dynamic and stochastic behavior of these systems. Therefore, simulation is widely applied for modeling and analysis of manufacturing systems (e.g., Bhaskaran 1998, Petrovic 2001). However, expensive model building is one of the main drawbacks of simulation (Law and Kelton 2001). High model building expenses may be especially preventive, if simulation is applied in preliminary studies of the system or as a supplementary tool. For instance, a simulation model can be used to examine analytical models under conditions not observable on the current system (Ignall et al. 1978).

Several approaches aimed to reduce expenses of the simulation model building have been proposed in the literature. Baker (1997) describes a methodology for incorporating classic operations research models into simulation models. Numerical algorithms are implemented using high-level programming languages. A simulation software user can insert these routines in the model and manipulate them through a user-friendly interface. Swaminathan et al. (1998) develop an agent-based simulation model building approach, where predefined agents are responsible for performing standard functions of supply chain management. A supply chain simulation model can be developed by assembling these agents. Son et al. (2000) describe an initiative by the National Institute of Standards and Technology to develop neural libraries of simulation model components in order to reduce simulation model building efforts. The standardized model components are used to develop a simulation package independent model. A package specific translator is used to generate an executable model. Werner and Weigert (2002) describe integration between a specific manufacturing and planning simulation model and an enterprise resource planning (ERP) system. The simulation model is continuously recreated using the most recent data retrieved from the ERP system.

There are a number of other works advocating the use of templates to improve model building efficiency (e.g., Pater and Teunisse 1997). However, the usage of templates or library objects still requires a substantial number of manual operations, especially, in modeling large scale systems having a large number of objects.

This paper describes a data driven simulation model building approach. The approach is designed for developing simulation models to be applied for preliminary analysis of multi-stage, multi-product manufacturing systems. It is aimed to reduce efforts associated with model building. For this purpose, manufacturing units are approximated by their generic representation which captures common functions of the units such as, handling of incoming product flows, flow transformation and handling of outgoing product flows. The manufacturing system is defined using a data model. The data model is assembled and subsequently standardized from several raw data sources. It contains information about both structure and operational characteristics of the system. A simulation model generator uses these data to automatically create a simulation model. The model is created on the basis of a predefined template.

*Proceedings 15th European Simulation Symposium*
*Alexander Verbraeck, Vlatka Hlupic (Eds.)*

*ISBN 3-936150-28-1 (book) / 3-936150-29-X (CD)*

The proposed model building approach is applied for testing scheduling algorithms at stamping plants of an automotive company. Modeling of each stamping plant includes dealing with several units such as external suppliers; blanking, pressing and assembly departments. Each plant processes up to a thousand products and their components using more than a hundred work stations. Data necessary for the model building are extracted from several different data sources such as relational data bases. The automated model building approach is attractive because, (a) the simulation model is to be used as the supplementary tool for testing of scheduling algorithms (i.e., only limited resources for simulation model building are provided), (b) there is a large number of similar objects in the system (e.g., at certain level of abstraction, processing of all products is almost identical) and, (c) all stamping plants are similar though their dimensions vary.

The main contributions of this research are as follows:

- Elaborating tools for data gathering from multiple sources;
- Developing a standardized data representation of multi-stage manufacturing systems;
- Separating the simulation model from its input data;
- Developing tools for automated model generation.

The rest of the paper is organized as follows. The following section describes the conceptual framework. It is followed by more detailed description of the proposed methodology and application examples.

## MODEL BUILDING APPROACH

The proposed simulation model building approach utilizes two main concepts: 1) separation between data and the model; and 2) a generic representation of manufacturing units. The main stages of the model building approach are shown in Figure 1.

Data necessary for the simulation model building are located in several different raw data sources such as, data bases and spreadsheets. A data converter is used to gather these data from all sources and to create a data model. The data model represents the data in a format suitable for generation and execution of a simulation model.

The data sources may have different formats, and definitions of data fields may differ among data sources. Therefore, a modeling taxonomy is used to establish standardized data definitions (Chandra et al. 2002). The taxonomy is a single and comprehensive source of information about characteristics of a general system. An ontology enables communications between different components of the particular system under consideration on the basis of the standardization provided by the taxonomy. In this case, the ontology defines mapping between the data sources and the data model. It defines both content and structure of data.

A model generator automatically creates a simulation model using the data provided in the data model. It assumes certain characteristics of the manufacturing system. Manufacturing units involved in the system are believed to have common functions. These common functions are handling of incoming and outgoing flows, flow transformation and control. A generic unit performing all these functions is constructed (Figure 2). Each manufacturing unit is approximated by its generic representation. The control function determines the way each generic function is performed at a particular unit. For instance, the handling of outgoing flows can be performed in either a push or pull manner. A network of the generic units represents the entire multi-stage manufacturing system.

The simulation model is generated on the basis of predefined template. The template does not contain any simulation objects. It only contains procedures for executing control of the generic functions and data declarations. The procedures have a uniform design. Different procedures can be developed to perform the same activity. Thus, different management policies can be analyzed.

Intermediate data are used to improve efficiency of data exchange between the data model and the simulation model.

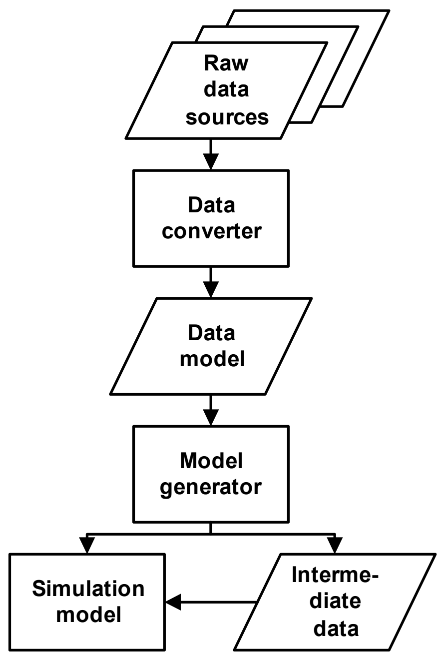


Figure 1: The Model Building Approach

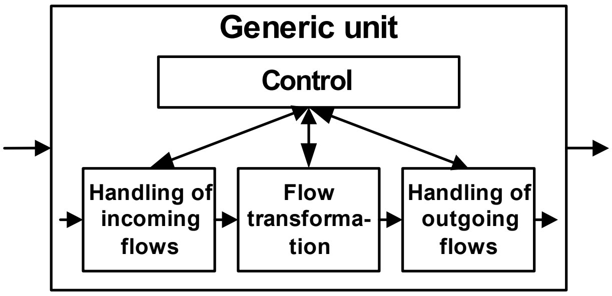


Figure 2: The Generic Unit

## MODEL BUILDING STAGES

### Data model

The data model organizes data describing the system in a manner suitable for the simulation model building and execution of the simulation model. These data describe structure of the system, properties of production units and products produced, and relationships of the system with its external environment including customers. For purposes of execution of the simulation model, structuring of the data should ensure quick access of necessary data items.

The data converter generating the data model is shown in Figure 3. The taxonomy contains a standardized description of concepts relevant to the system. It is not built specifically for the simulation model building, but is a part of more general enterprise wide initiative for data standardization. The taxonomy facilitates common understanding for terms such as products, resources, processing time, etc.

The data conversion process is illustrated by an example. The raw data source contains data fields characterizing throughput for each resource in items per hour and a corresponding efficiency measurement in percents. The converter uses these data fields to determine processing time for each resource in hours per item (this values is used by the simulation model) and places the derived data item in the appropriate position of the data model. Another example is translation of the term Work Center used in one of the raw data sources. The converter identifies this term with the taxonomical term Resource, which is also defined in the same way in the simulation data model. If a product is produced internally, it should have at least one resource assigned. But the generic representation of units requires a resource assigned and products purchased from external suppliers. Therefore, the converter assigns a dummy resource to the external products. These and similar conversion rules are described in the ontology.

The data model consists of multiple tables containing information about structure and operational characteristics of the system. The structure of the system is described by bill of materials, etc. The structural information is also represented using several specialized tables, which are designed to facilitate data retrieval by the simulation model. The operational characteristics describe processing time, setup time, transportation time. The data model is implemented as a Microsoft Excel workbook. Table 1 lists tables included in the data model. The concept of product and unit pair is introduced to make distinction between the same products processed at different units. Time parameters are specified using a string describing a probability distribution.

The elaborated data model representation allows describing a wide range of manufacturing networks. The main characteristics of these networks are as follows. A product can be produced at several units and it can be a component of several products produced at different units. A resource belongs to one particular unit (as specified in Table ResourceUnit). It has finite capacity. Several products may share the same resource, and a product can be produced by using alternative resources as specified in Table PairResource.

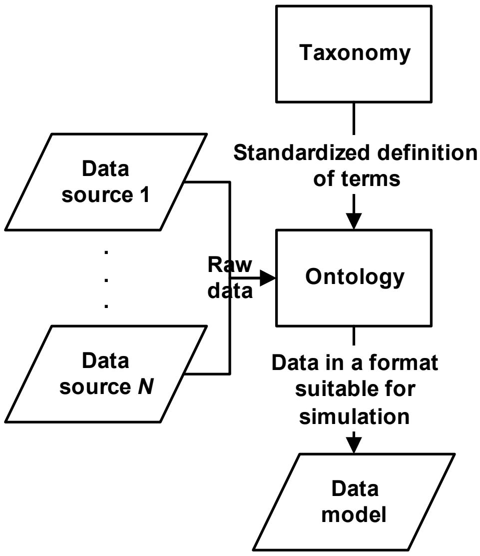


Figure 3: The Data Converter

Table 1: Tables of the Data Model. Type S Refers to Structural Data and Type O Refers to Operational Data.

| Table | Type | Description |
|---|---|---|
| Definitions | S | Dimensional data (e.g. number of products) and modeling control data (e.g. number of replications) |
| Demand | O | Customer demand per week |
| Schedule 1 | O | Scheduled order sizes |
| Schedule 2 | O | Scheduled resource assignments |
| UnitsProducts | S | Shows products produced by each unit |

| Table | Type | Description |
|---|---|---|
| Pairs | S | Defines pairs |
| PairDestinations | S | Defines possible destinations for a product from the pair |
| ResourceUnit | S | Shows resources for each unit |
| PairResource | S | Shows which resources can be used to process a product from the pair |
| BOM 1 | S | Bill of materials, indicates components of each product by component number |
| BOM 2 | S | Bill of materials, indicates items of each component needed |
| SetupTime | O | Setup time for products according to resource used |
| ProcessingTime | O | Assembly time for products according to resource used |
| TransTime | O | Transportation time for products according to destination |
| ResourceFailure 1 | O | Time between two consecutive resource failures |
| ResourceFailure 2 | O | Resource downtime duration |

**Simulation model**

The simulation model is automatically created by the model generator. The model generator creates one submodel for each generic unit and one submodel to represent external customers. The simulation model is generated in the ARENA simulation modeling environment (Rockwell Software 2001). A generated submodel representing the generic unit is shown in Figure 4. The representation of the manufacturing system consists of several such submodels.

Block 1 at the begging of each period (week) generates an entity representing a production order for each product produced by a particular unit. Block 2 assigns values of identification attributes to the entities. Block 3 reads data from the production schedule provided in order to determine the batch size and the resource to be used. The block has capabilities to change initial resource assignments according to current circumstances. The scheduled resource assignments can be changed, if the schedule does not contain any assignments. One entity represents the entire batch of products. After leaving Block 3, the entity carries information about the product it represents, the batch size and the resource assigned. The entities are held in Block 4 until the assigned resource and all components of the product become available. Each product is held in its own queue, and the holding condition is also product specific. These queues and holding conditions are organized using the set of queues, and the set of expressions option, respectively. The production setup process is represented by Block 5. The setup process requests a product specific resource according to the assignment. The processing time depends upon the product and the resource used. Block 6 is used for additional checking for availability of components before the final assembly is started. Block 7 retrieves the components from the inventory. The inventory is represented using a multi-dimensional array. The assembly process is represented by Block 8. It requests a product specific resource according to the assignment. The processing time depends upon the product and the resource used. Block 9 checks whether to change the scheduled resource assignments. Block 10 splits the production batch in transportation batches (products can be sent to different parent units). The production batch is split in as many transportation batches as the number of parent units. Block 11 determines the size of each transportation batch according to its intended destination. The transportation process is represented by Block 12. The transportation time depends upon the product and the destination. Block 13 represents receiving of components from other units. Block 14 updates inventory data for the components. Block 15 checks whether or not to change the scheduled resource assignments.

The generated model is the conventional ARENA model. A user can edit the model, use the standard output reporting features and perform other manipulations.

At the beginning of simulation, modeling data from the data model are loaded in the simulation model. Before loading, the intermediate data have been created by converting the data model tables from the Excel format into the text format because ARENA reads text files much faster than Microsoft Excel files. Some of the data tables are loaded into ARENA arrays for access by ARENA objects, while some others are loaded in Visual Basic arrays for access by control functions.

All VBA blocks invoke a main Visual Basic procedure. An entity attribute characterizing the function to be performed is assigned to the entity before it enters a VBA block. The main procedure reads attributes of the entity to determine a specific procedure to be called and parameters of the specific procedure. The specific procedures read data from the production schedule, reassign resources, update inventory data, check material availability, determine transportation route for products, etc. All procedures are part of the model template.

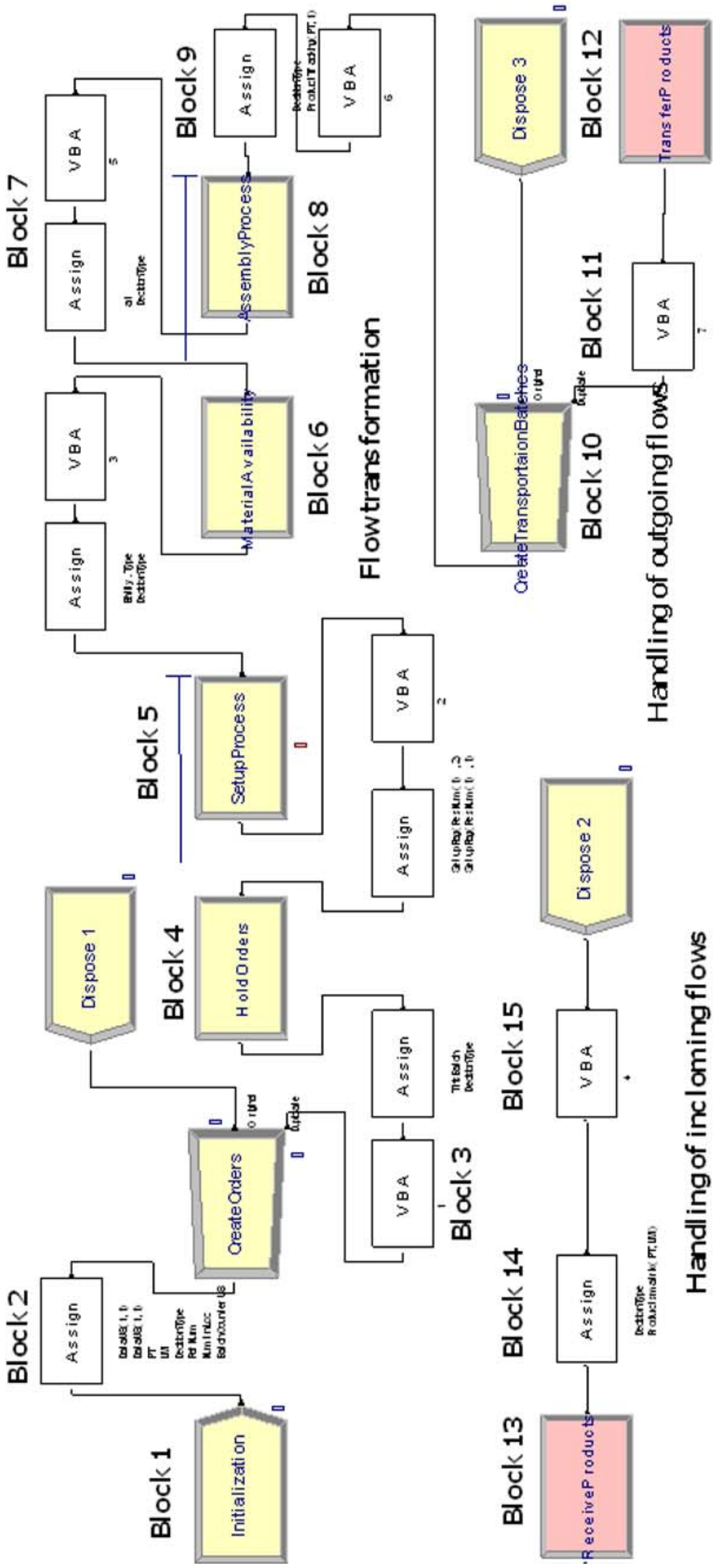


Figure 4: Representation of the Generic Supply Chain (ARENA implementation).

VisuaBasic. It creates ARENA objects using the ActiveX technology. Actually, the same data model can be used to create a simulation model in other simulation modeling environment supporting the ActiveX technology. A similar model generator has been developed for creating models in ProMODEL. However, ARENA appears to be more flexible mainly because of better support for integration with high level programming languages, easier generation of animation, higher level of openness to user editing, and more flexible input and output features.

The simulation model needs to be regenerated, if the structural data tables have been changed. Changes in the operational data tables can be captured just by updating the intermediate data.

Development of the model generator can also be a labor-intensive task. However, this model generator has been derived from a more general supply chain simulation model generator created by the authors. Modifications are introduced to represent some specific properties of the particular manufacturing system. Additionally, the usage of the model generator effectively eliminates syntactic and logical errors, which are likely to occur because of a large number of objects.

## APPLICATION

The elaborated model building approach and automatically generated simulation models are used in experimental studies of a multi-stage manufacturing system in order to evaluate applicability of the model building approach and to improve performance of the system analyzed.

### Case Description

Case studies using the elaborated model building approach are conducted based on a modeling problem experienced by an automotive company. The manufacturing system considered consists of raw material suppliers and a stamping plant (a similar steel processing supply chain at General Motors has been analyzed by Bhaskaran (1999)). The system is expected to meet strict delivery time requirements. Meeting of these requirements often force the plant to use a premium cost transportation mode. The modeling objective is to determine whether additional transportation costs are caused by:

- insufficiently coordinated deliveries of raw materials;
- production scheduling inefficiencies;
- faults in operations.

The manufacturing system produces approximately 300 end products using approximately 1000 components. Manufacturing operations are performed using approximately 100 work centers. The end-products are delivered to about 30 customers, which are assembly plants and repair centers. Each customer places orders for multiple products.

The stamping plant consists of blanking, pressing and assembly departments. The blanking department cuts the raw steel into rectangular pieces. Work centers at this department are relatively flexible to process different products and setup times are insignificant. The pressing department stamps the blanks into parts. Work centers at the pressing department are partially specialized. There are substantial setup times. Welding and other operations are performed on stamped parts at the metal assembly department. Works centers at the assembly department are specialized where setup times are smaller than at the pressing department. Transportation times within the plant are assumed to be insignificant.

Production is initiated according to a production schedule. The production schedule is elaborated according to weekly customer demand. It specifies the quantity of products to be produced and resources (i.e., work centers) to be used in production of these products. The production schedule is implemented in the rolling horizon environment. The resource assignments can be dynamically changed to adjust for the actual state of the system.

Majority of costs in the system are fixed. Variable costs are the inventory holding cost and the transportation cost. The transportation cost consists of the cost for a standard mode of transportation and the cost for a premium mode of transportation. The standard mode of transportation is used for on time deliveries. The premium mode of transportation is used, if deliveries of ordered products are delayed.

The stochastic factors in the system are setup times, processing times and resource failures. Additionally, external demand used to elaborate the production schedule is stochastic. However, the demand for the current production period is fairy stable.

The company operates a relatively large number of similar manufacturing systems. In this paper, modeling is conducted only for one of them. However, the same data model populated with appropriate data and the simulation model generator can be used to automatically generate simulation models for other related manufacturing systems.

### Experimental Design

The third modeling objective on faults in operations is addressed. Particularly, the impact of setup time uncertainty on the production performance is analyzed. The setup time uncertainty is caused by a number of factors many of which are supposed to be avoidable. Three levels of the setup time uncertainty are considered. These levels include deterministic setup time, the standard deviation of the setup time equal to 10% of the average setup time and the standard deviation of the setup time equal to 20% of the average setup time. In the cases with the stochastic setup time, it is modeled using the lognormal distribution.

The systems performance is measured by waiting time of customer orders. If the waiting time is zero then the customer orders are satisfied without relying on the premium transportation. However, if the waiting time is larger than zero, then the premium transportation is to be used. In the real system, the waiting time is not allowed to exceed certain threshold even when the premium transportation cannot assure timely deliveries. This aspect is currently ignored. The waiting time is a proxy measure for the premium transportation cost.

The system is modeled for one year and five replications are conducted for each level of the setup time variability.

The model does not represent a number of constraints and operations of the real system. For instance, the model does not represent the restriction that only a limited number of setups can be performed simultaneously. Impact of these constraints and operations are assumed to be insignificant. Additionally, many decisions such as prioritizing deliveries among customers are done in a non-formal manner, thus, making these decisions difficult to model and validate.

**Experimental Results**

Figure 5 shows the average customer waiting time according to the setup time variability. The waiting time is expressed relative to the waiting time in the case of the deterministic setup time. The results indicate that the customer waiting time substantially depends upon the setup time variability. The relationship is approximately linear. However, the waiting time is larger than zero (i.e., the premium transportation mode is to be used) also for the case with the deterministic setup time. Therefore, improving the setup time variability is not the only solution to the transportation cost reduction. There are multiple reasons for the setup time variability. The variability is caused by workforce resource limitations at the work floor level, qualification of workforce, quality of raw materials and precision of manufacturing tools. Dealing with these issues may require substantial organizational changes.

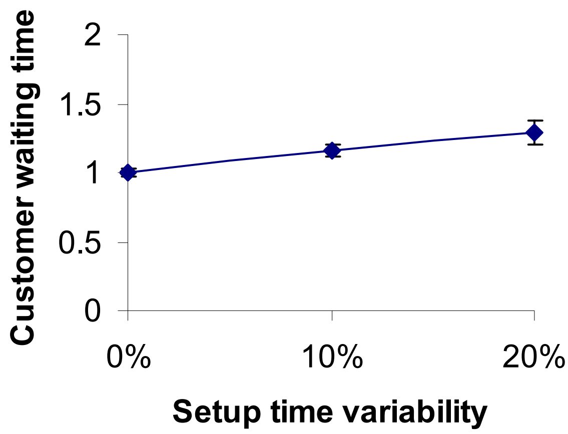


Figure 5: The Average Relative Customer Waiting Time According to the Setup Time Variability

Highly variable resource utilization is another problem faced by the manufacturer (Figure 6). Some of the resources are nearly overloaded, while others have low utilization rates. Achieving a more uniform distribution of the workload among resources would also facilitate reduction of the customer waiting time. However, this is constrained by inflexibility of resources, which are capable of processing only a limited number of products.

**CONCLUSION**

The automated simulation model building approach

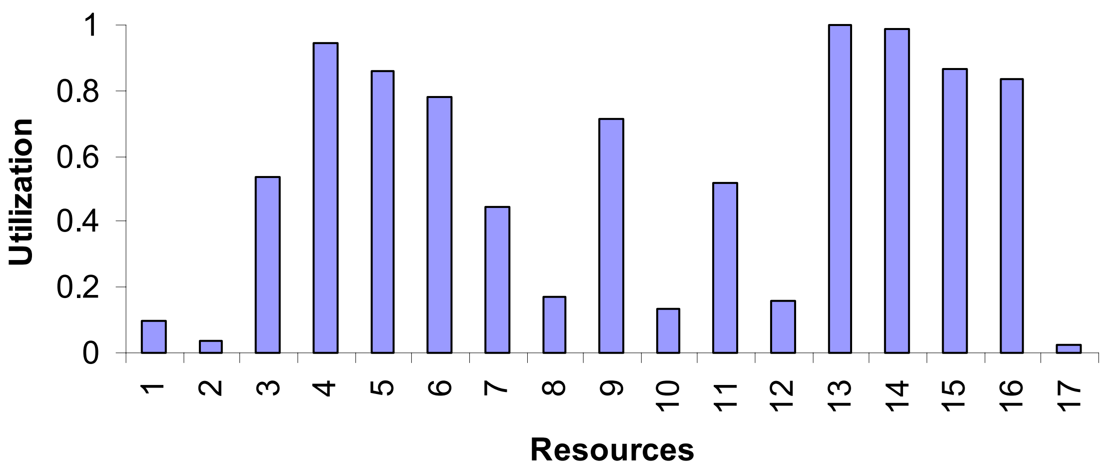


Figure 6: Average Utilization of Resources at One of the Manufacturing Units

has been elaborated. The approach is aimed at (a) reducing simulation model building efforts, (b) reducing model building errors, and (c) providing reusability. The simulation model is generated according to the problem definition provided by the data model which describes a multi-stage manufacturing system in the standardized manner. It is developed by assembling and standardizing raw data characterizing the system from multiple data sources. The data are standardized using the taxonomy and ontology concepts. Manufacturing units in the simulation model are represented using their generic approximations which describes the common functions of manufacturing units. The generated simulation model is open for customization.

Advantages of the elaborated model building technology are:

- Model building efforts are reduced by using the generic approximation of manufacturing units;
- The number of modeling errors is reduced because the large scale structure of the system is generated automatically instead of manual input;
- Editing of the simulation model is made more efficient because changes can be introduce by simple updating of the data model;
- The modeling process can be repeated for multiple similar manufacturing facilities without substantial model building efforts;
- Systematization of the problem analysis by establishing standardized definitions for subjects involved in the system.

The proposed model building approach and the generated simulation model have been applied to study the multi-stage manufacturing supply chain. Objectives of the studies are to identify factors reducing efficiency of manufacturing operations and to test alternative production schedules. The experimental results suggest that the setup time variability has substantial adverse impact on efficiency of the manufacturing system considered.

There are substantial problems associated with validation of the model. Two of the main obstacles are that the actual system heavily relies on judgmental decisions made by production managers and a lack of reliable data quantifying operational properties of the actual system. Accumulation of data needed for a thorough validation would incur substantial additional expenses. Currently, the discussion of the results with area specialists is the main validation approach. Additionally, validation should be performed with respect to the level of approximation provided by the model (an account on validation of models at different levels of abstractions can be found in Persson (2002)).

## REFERENCES


Baker, G.S. 1997. "Taking The Work Out Of Simulation Modeling: An Application Of Technology Integration." In *Proceedings of the 1997 Winter Simulation Conference* (Atlanta, GA, Dec. 7-10), 1345-1351.

Bhaskaran, S. 1998. "Simulation analysis of a manufacturing supply chain." *Decision Sciences* 29, 3, 633.

Chandra, C.; J. Grabis; R. Marukyan; and A. Tumanyan. 2002. Supply chain taxonomy: Development and applications, *Research and Application Congress, SAP Institute for Innovation and Development* (Tampa, FL, Feb. 23-25), 2002.

Ignall, R.G.; P. Kolesar; and W. Walker. 1978. "Using simulation to develop and validate analytical models." *Operations Research* 28, 237-253.

Law, A.M.; and W.D. Kelton. 2000. *Simulation Modeling and Analysis*. McGraw-Hill, New York.

Pater, A.J.G.; and M.J.G. Teunisse. 1997. "Use of a template-based methodology in the simulation of a new cargo track from Rotterdam harbor to Germany". In *Proceedings of the 1997 Winter Simulation Conference* (Atlanta, GA, Dec. 7-10). 1176-1180.

Persson, J. F. 2002. "The impact of different levels of detail in manufacturing systems simulation models." *Robotics and Computer-Integrated Manufacturing* 18, 3-4, 319-325.

Petrovic, D. 2001. "Simulation of supply chain behaviour and performance in an uncertain environment." *International Journal of Production Economics* 71, 1-3, 429-438.

Rockwell Software. 2001. ARENA: User's Guide. Sewickley: Rockwell Software Inc.

Son, Y.J.; A.T. Jones; and R.A. Wysk. 2000. "Automatic Generation of Simulation Models From Neutral Libraries: An Example." In *Proceedings of the 2000 Winter Simulation Conference*, 1158-1567.

Swaminathan, J. M.; S. F. Smith; and N. M. Sadeh. 1998. "Modeling supply chain dynamics: A multiagent approach." *Decision Sciences* 29, 3, 607.

Werner, S; and G. Weigert. 2002. "Process Accompanying Simulation – A General Approach For The Continuous Optimization Of Manufacturing Schedules In Electronics Production." In *Proceedings of the 2002 Winter Simulation Conference*, 1903-1908.


## AUTHORS BIOGRAPHY

**CHARU CHANDRA** is an Associate Professor in Industrial and Manufacturing Systems Engineering at the University of Michigan-Dearborn, U.S.A. Prior to this Charu was a Post Doctoral Fellow at Los Alamos National Laboratory, Los Alamos, New Mexico, and at the University of Minnesota, Minneapolis, U.S.A. He has worked in the industry as Information Technology Manager and Systems Analyst. He is involved in research in Supply Chain Management, and Enterprise Integration issues in large complex systems.

Specifically, his research focuses on studying complex systems with the aim of developing cooperative models to represent coordination and integration in an enterprise. He has published several papers and book chapters in leading research publications, in areas of supply chain management, enterprise modeling, inventory management, and group technology. He teaches courses in Information Technology, Operations Research and Supply Chain Management. His Ph. D. degree is in Industrial Engineering and Operations Research from the Arizona State University. He is a member of Institute of Industrial Engineers, Institute of Operations Research and Management Sciences, Decision Science Institute, Production and Operations Management Society, American Association of Artificial Intelligence, and Association of Information Systems.

**JĀNIS GRABIS** obtained his Ph. D. degree in Information Technology from the Riga Technical University in 2001. He spent two years with the University of Michigan-Dearborn and currently works in the Riga Technical University. His main research interests are supply chain management, simulation, forecasting and software project management.