Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A.A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

COMBINING DATA FARMING AND DATA ENVELOPMENT ANALYSIS FOR MEASURING PRODUCTIVE EFFICIENCY IN MANUFACTURING SIMULATIONS

Niclas Feldkamp
Soeren Bergmann
Steffen Strassburger

Department for Industrial Information Systems

Ilmenau University of Technology
P.O. Box 100 565
Ilmenau, 98693, GERMANY

Erik Borsch
Magnus Richter
Rainer Souren

Department for Sustainable Production and Logistics Management
Ilmenau University of Technology
P.O. Box 100 565
Ilmenau, 98693, GERMANY

ABSTRACT

Discrete event simulation is an established methodology for investigating the dynamic behavior of complex manufacturing and logistics systems. In addition to traditional simulation studies, the concept of data farming and knowledge discovery in simulation data is a current research topic that consist of broad scale experimentation and data mining assisted analysis of massive simulation output data. While most of the current research aims to investigate key drivers of production performance, in this paper we propose a methodology for investigating productive efficiency. We therefore developed a concept of combining our existing approach of data farming and visual analytics with data envelopment analysis (DEA), which is used to investigate efficiency in operations research and economics. With this combination of concepts, we are not only able to determine key factors and interactions that drive productive efficiency in the modeled manufacturing system, but also to identify the most productive settings.

1 INTRODUCTION

Data farming and knowledge discovery in simulation are ongoing research matters in current simulation methodology. The concept of data farming encourages large scale and data intensive simulation experiments in order to cover a broad and profound bandwidth of possible system behavior. In order to investigate massive amounts of simulation data, algorithmically supported analysis like data-mining-methods are applicable. Still, the development of new ways for analyzing large amounts of data farming output data is an important and current research issue. Therefore, this paper investigates the application of data envelopment analysis (DEA) in combination with data farming in order to investigate system behavior in terms of productive efficiency.

Simulation models have very rarely been used in combination with DEA so far. Publications are either using DEA in combination with simulation-based optimization algorithms (Aminuddin and Ismail 2016; Fazli et al. 2012; McMullen and Frazier 1998; Weng et al. 2011) or to simulate missing data for existing real-world data sets in order to be able to applicate DEA (Mahfouz and Arisha 2015; Min and Park 2008). In all of the work mentioned above, DEA method is applied on very few targeted, goal-driven experiments, while our approach proposes large-scale experimentation and therefore in-depth analysis of productive efficiency and its relations and interactions to the simulation model's factors. Therefore we are able to investigate efficiency and drivers of efficiency of the modeled system in an in-depth manner. In this context, the efficiency scores generated with DEA may either be used as an additional performance indicator or, depending on the inputs and outputs taken into account, as an overall measure aggregating commonplace performance criteria, like, e.g., cycle time or utilization. This approach enables decision makers in the planning phase of a manufacturing system to get a sense of efficient use of resources while maintaining system performance or fine-tune existing systems in terms of efficiency.

The remainder of the paper is structured as follows. In section 2, we introduce the related work regarding data farming and knowledge discovery in simulation data, and we give a short introduction to the DEA method. In section 3, we present our concept of how to integrate DEA and data farming, followed by a demonstration by means of a case study in section 4. In section 5, we give some concluding remarks and a discussion of possible future work.

2 RELATED WORK

2.1 Data Farming and Knowledge Discovery in Simulation Data

The concept of data farming refers to a methodology for using a simulation model as a data generator. By using efficient experimental design alongside high performance computing, one is able to maximize data yield and corresponding information gain (Elmegreen et al. 2014; Horne and Meyer 2005). The farming metaphor describes how the data output can be maximized by experimental designs like a farmer that cultivates his land to maximize his crop yield (Sanchez 2014). This is in particular made possible by new approaches in the design of experiments that are highly efficient and manage a balance between broad scale parameter combinations and manageable data volume (Kleijnen et al. 2005).

Feldkamp et al. (2015a) developed a method for using data farming in manufacturing simulations for finding hidden patterns and relations in large quantities of simulation output based on broad scale experimentation and visual aided analysis called knowledge discovery in simulation data. The first step after the farming of data is to preprocess the simulation output data with data mining algorithms, then knowledge can be gained through visualizations of data mining results combined with visual representations of the input/output relations in the target data. While visualizations are commonly applied in almost any discrete event simulation study in terms of animation, time-plots or graphs of confidence intervals of certain performance indicators, this approach presented in previous papers goes beyond commonly applied techniques by making visualizations the central anchor point of an iterative and interactively driven exploration of simulation data analysis (Feldkamp et al. 2015b). This is based on the research area called visual analytics (VA). VA is defined as "an iterative process that involves information gathering, data preprocessing, knowledge representation, interaction and decision making" (Keim et al. 2008). By combining automated data analysis and interactive visualizations, it also combines the strengths of both machine and human capabilities. On the one hand, patterns from large amounts of data can be extracted and processed through data mining with statistical and mathematical models. This is commonly referred to as knowledge discovery in databases (Fayyad et al. 1996). On the other hand, visualizations of the processed data can be explored by making use of the human capabilities to perceive, relate, and recognize visual patterns and draw conclusions, encouraged by a high degree of user interaction (Thomas and Cook 2005). For the algorithmically supported side of the analysis, we already showed the application of various methods onto large amounts of simulation data, e.g., clustering (Feldkamp et al. 2015a), decision trees or Taguchi's loss function for robustness analysis (Feldkamp et al. 2017). In this paper, we incorporate a first approach to use DEA for the analysis of productive efficiency. Therefore, we give a brief review on the background of this method in the next section.

2.2 Data Envelopment Analysis (DEA)

DEA (Charnes et al. 1978) is an instrument to determine the relative efficiency of decision making units (DMUs) using multiple inputs i and outputs j ($i = 1, \ldots m$; $j = m+1, \ldots, m+n$; $m, n \geq 2$) (Dyson et al. 2001). Misleadingly, the term *decision* making unit not necessarily implies a real decision-making authority of the productive units under evaluation. In fact, within DEA, quite different kinds of productive units can serve as DMUs, even if they only *execute* managerial decisions. Consequently, not only companies or facilities can be used as DMUs but also working stations, machines or employees. Popular (non-industrial) examples of DMUs used for DEA-based efficiency measurement are universities and hospitals as well as the corresponding subunits like, e.g., research projects and medical wards (for a bibliography of applications see (Seiford 1997)).

Though, DEA is mainly used for evaluation of real life production systems and activities it should be considered more comprehensively as a general approach for solving all different kinds of multi criteria decision making (MCDM) problems. Generally, both inputs and outputs of DMUs can either be seen as *quantities of objects*, like, e.g., amounts of materials being used or units of products being manufactured, or they serve as *proxy attributes*, representing managerial objectives, like, e.g., cycle times or utilization. The appropriate specification of DMUs, as well as the selection of inputs and outputs, strongly depends on the analysts' aims and scopes and the specific circumstances of the system under consideration. In the paper at hand, simulation runs, representing different settings of a virtual manufacturing systems, are used as DMUs.

In contrast to the fundamental economic efficiency concept proposed by (Pareto 1909) and (Koopmans 1951), DEA is *not* limited to the dichotomous distinction between "efficient" and "inefficient" activities, but additionally provides information on the *degrees* of inefficiency. This allows distinct evaluations of the DMUs under consideration even in cases where multiple performance criteria have to be captured and aggregated, respectively, without knowing their relative importance. This might be the case when, e. g., a manager pursues different objectives at a time or when prices of inputs or outputs are unknown or unavailable to the decision maker. Then, DEA can be a powerful means of compensation for the lack of such information.

For each single DMU Ω ($\Omega = 1, \ldots, O$) DEA calculates a one-dimensional efficiency score ϕ^{Ω} using an *individual* set of weights c_i (for the inputs i) and e_j (for the outputs j) that maximizes the ratio between weighted output and weighted input (Charnes et al. 1978).

$$\max \qquad \phi^* = \frac{\sum_{j=m+1}^{m+n} e_j \cdot y_j^*}{\sum_{i=1}^{m} c_i \cdot x_i^*} \tag{1}$$

$$\text{so that} \qquad \frac{\sum_{j=m+1}^{m+n} e_j \cdot y_j^{\Omega}}{\sum_{i=1}^{m} c_i \cdot x_i^{\Omega}} \leq 1 \qquad \Omega = 1,\ldots,O \tag{2}$$

$$\text{and} \qquad c_i, e_j \geq 0 \tag{3}$$

In simple terms, the algorithm underlying DEA choses weights for the inputs and outputs of the DMU under evaluation (marked as *) individually, presenting the DMU "in the best light" (see for a more detailed description of this procedure (Cooper et al. 2006)). Taken together, the terms (1), (2) and (3) illustrate the formal structure of the standard DEA model. The weights used for input and output quantities of the DMU under consideration may not lead to efficiency scores greater 1 when applied to the corresponding quantities of the remaining O–1 DMUs within the reference set (see (2)). The weights can be interpreted as (artificial) unit prices of the inputs and outputs, as they reflect the relative importance of the factors and products within the efficiency analysis. If all inputs and outputs are classified as goods (desirable objects), the prices c_i and e_j are positive (see (3)). Moreover the DEA model can be modified by introducing weight restrictions to capture further characteristics of the real production system under evaluation, such as preferences for certain products (e.g. "weight of *throughput A* > weight of *throughput B*"). Furthermore, for each DMU under evaluation DEA reveals efficient benchmarks (or combinations of such benchmarks), which can be used as "role models" for a better performance.

3 APPLYING THE DEA METHODOLOGY ON SIMULATION DATA

While usually simulation studies investigate "good system performance" in terms of throughput, cycle time, utilization or robustness, productive efficiency ϕ^*, as defined within DEA (see again (1)), is often neglected, but equally important. This requires the identification of system configurations where input resources are used efficiently, which is not necessarily the case when looking only at other performance indicators. An efficient system configuration does not necessarily imply a good system performance et vice versa. On one hand, if, e. g., the system's throughput is unacceptably low, it still might be an efficient solution, and on the other hand, a high product throughput may be due to an unreasonably high input of resources.

Therefore, the investigation of productive efficiency in a massive simulation output database acquired through data farming must include several consecutive steps. First, DEA has to be applied on the simulation output data after experiments have been fully conducted. Each simulation experiment represents a decision making unit for the DEA model. The DEA output is obviously very easily definable as it corresponds to the specific values of the output parameters of the simulation model, preferably the throughput of certain product types. The DEA input generally corresponds to the input parameters of the simulation model i.e. the factors of the experiment design. However, when connecting DEA input to the simulation model's input, one should be cautious since not every simulation input parameter can be considered as a production input resource from an business economics point of view.

After DEA is applied, it ranks every simulation experiment with an efficiency score, so experiments can be sorted or filtered accordingly. In the next step, the investigation of those experiments enables the exploration of the performance of experiments and their relation to the efficiency. Furthermore, in complex models in the context of production and logistic, performance is multidimensional, which means not only throughput is relevant, but also subsequent performance measures like e.g. cycle times or utilizations of system elements need to be in a desired range. Since performance measures might possibly be in conflict with each other, we propose an interactive visually guided analysis that allows some sort of user-defined sensitivity, what we already demonstrated in our recent work (Feldkamp et al. 2015a, 2015b). This also includes the investigation of corresponding input parameters (experiment factors) in terms of what input parameter values account for good and efficient system performance, and what distinguishes efficient from inefficient experiments.

Furthermore, note that we propose the use of crossed experiment designs. This arises from the assumption that not all of the factors that are controllable in the simulation model, are in fact controllable in the corresponding real world model (Sanchez 2000). For example, this is typically the case in the automotive industry, where productions are highly customizable and fluctuations in customer demand can lead to variation in the mixture of jobs that need to be dispatched in the system. For this purpose, a separation of factors into two categories called decision and noise is useful. When crossing independent experimental designs for both of those categories, we are able to investigate how each system configuration holds up against variations in the noise factors.

4 AN EXPLORATORY CASE STUDY

4.1 Model Description and Design of Experiments

For a first proof of concept, we developed a discrete event simulation model of a typical assembly line. This model was implemented in Siemens Plant Simulation, as shown in Figure 1. In this model, five different part types are loaded onto carriers that are transported on a conveyor. Parts are both automatically processed on assembly stations and manually treated on up to five workplaces. Before getting unloaded from the carrier, parts go through a manual quality assurance (QA). If they pass, they can leave the system or are otherwise returned to the manual workplaces.

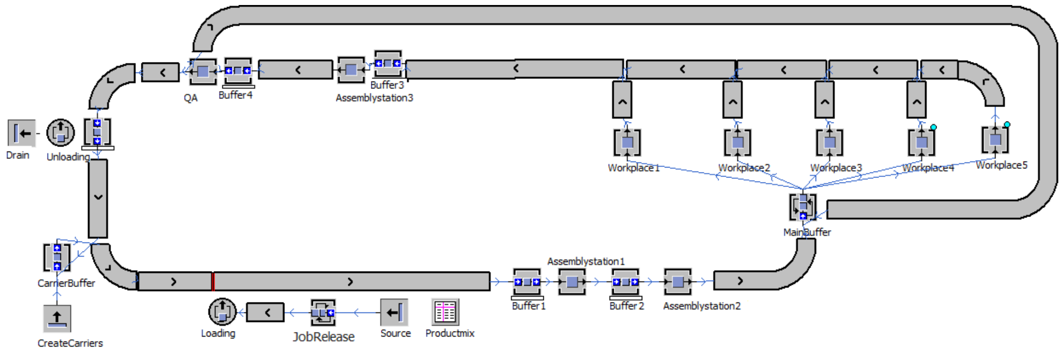

Figure 1: 2D view of the assembly line model.

The mixture of parts can vary, but arriving parts are kept in a buffer until they are cleared to get mounted on a free carrier. Some stochastic effects are implemented regarding machine reliability and a small proportion of parts that fail the quality assurance. For the experimental design of the decision factors, we used a *nearly orthogonal latin hypercube (NOLH)*, which is commonly used in data farming and is much more efficient than a full factorial n^k-design (Vieira et al. 2011). Table 1 shows the decision factors in the simulation model. Note that the process times for workplaces and quality assurance station are stochastic and are normally distributed. We choose to only alter the variance of the distribution in a small range over the experiments because changing the distribution possibly would have impacted the results in such a strong way that it might superimpose any other interesting effects in our case study model and therefore would lower its exemplary purpose.

Table 1: Decision factors for the simulation experiments.

Factor name	**Scale**	**Description**	**Margins**
LoadingTime	Continuous	Duration for mounting parts	10-60s
UnloadingTime	Continuous	Duration for unmounting parts	10-60s
InterArrivalTime	Continuous	Interarrival time of jobs	100-300s
SortingStrategy	Categorical	Sorting strategy for jobs {fifo, lot size of: 5/10/unlimited}	1-4
NumberOfWorkplaces	Discrete	Number of manual assembly work places	1-5
NumberOfCarriers	Discrete	Number of carriers	1-100
MainSorterCap	Discrete	Capacity for main sorting buffer	1-100
WP_ProcTimeVar	Continuous	Allowed tolerance for workplace process time	100-300s
QA_ProcTimeVar	Continuous	Allowed tolerance for QA process time	100-300s
ConveyorSpeed	Continuous	Transportation speed of the conveyor	1-5 m/s

The resulting design has 512 design points, which were then crossed with 40 randomly distributed configurations for the product mixture resulting in 20.480 simulation runs. The final experiment design has been split into multiple files in order to be distributed onto ten machines. Result data was written into flat CSV-files and collected and aggregated through a dedicated database. After the experimentation was conducted and all result data was collected, we performed the DEA. The first step here is to choose suitable inputs from the list of experimental factors. As mentioned in Section 3, not all factors automatically qualify for usage as DEA input, because they have to represent some sort of production resource. For this purpose, we choose the two discrete factors *NumberOfWorkplaces, NumberOfCarriers* as they are integers that are countable and straightforwardly interpretable as resources. The capacity of the main sorter buffer is also taken into account since a buffer requires more spatial resources the bigger it is. Finally loading and unloading times for mounting parts on the carriers are taken into account because by increasing resources in terms of a countable workforce, we are able to quicker load and release the parts on and off the carriers. However, other interpretations of the model's factors are possible. Therefore, what factors qualify for DEA input is up to debate and depends highly on the simulation model and additional domain knowledge of the corresponding real world system is needed. Furthermore, we choose that in our model the DEA output corresponds to the throughput of each experiment. In fact, we modeled the throughput of each individual part type as an independent DEA output, since DEA allows the modelling of multidimensional input/output-relations. Theoretically, DEA allows for a big number of inputs (m) and outputs (n) as long as the equation $O \geq 2 \cdot m \cdot n$ is fulfilled (Dyson et al. 2001). The final DEA model is shown in Figure 2.

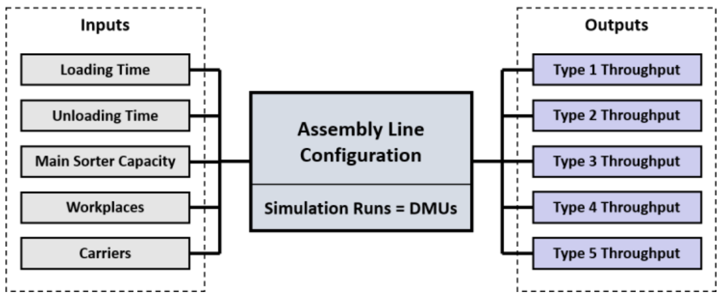

Figure 2: Input/output sets of the DEA model.

Each parameter setting represents a specific combination of input/output quantities and therefore serves as a distinct DMU (= unit of comparison). Due to the fact that different parameter settings (= combinations of input quantities) lead to different throughputs for each product type, the question arises which DMUs are overall efficient. The identification of an overall (or unambiguously) best DMU is particularly demanding, the more inputs and outputs are taken into account. If, e. g., setting *X* generates more *Throughput of Type 1, 2, 3 and 4* than setting *Y*, but – at the same time – has a lower *Throughput of Type 5* than setting *Y*, setting *X* cannot clearly be declared as "favourable", "dominant" or "better". To prevent a certain DMU from being declared as efficient due to only a single superior object quantity, weight restrictions for inputs and outputs can be applied. For reasons of simplicity this was skipped here. Nonetheless, for DEA applications on real data, weight restrictions can be a powerful means of matching the DEA model to the real world and, thus, to obtain more appropriate results. Particularly when conducted for supporting real life managerial decision making blind faith in quantitative DEA results without reflecting the underlying set of weights can be fatal, the more so as multiple feasible solutions might co-exist, each of which representing a different corporate target system.

4.2 Discussion and Results

The calculations of the input oriented efficiency scores of the 512 DMUs Ω (system configurations) in 40 different settings k (product mixes) were carried out with the software "MaxDEA". The calculation time of approximately seven seconds per setting is negligibly small. Because we are using a crossed design, we can calculate efficiency scores for every combination of decision factors with each of the product mixes. This investigation aims at examining the robustness of the efficiency of different system configurations against variation in the product mixture. Therefore, we calculated the mean $\bar{\Phi}$ and standard deviation s_Ω of the efficiency values Φ_k^Ω (one observation for each of the 40 product mixes), which serves as a measure for the robustness in the following meaning: the lower s_Ω, the less sensitive the system configuration reacts to changes in the product mixture and the more robust it is. Φ_k^Ω represents per se a measure for the performance of a system configuration. For that reason, the average efficiency $\bar{\Phi}^\Omega$ of a DMU Ω throughout all 40 settings might be regarded as an additional performance criterion when assessing a system configuration. This is shown in Figure 3. An average efficiency of 69,91 %, a range of efficiency values of $\left[\bar{\Phi}^{\Omega,\max}, \bar{\Phi}^{\Omega,\min}\right] = [100\%, 26{,}22\%]$ and the fact that less than 5 % of the system configurations are fully efficient according to DEA, demonstrate that the model is capable to discriminate clearly between the

DMUs. Altogether 24 system configurations are efficient according to the chosen model and simultaneously exhibit $s_{\Omega} = 0$. These system configurations are considered robust against variations of the product mixture and moreover efficient in terms of input and output factors (e.g. $\Omega = 14, 23, 26$).

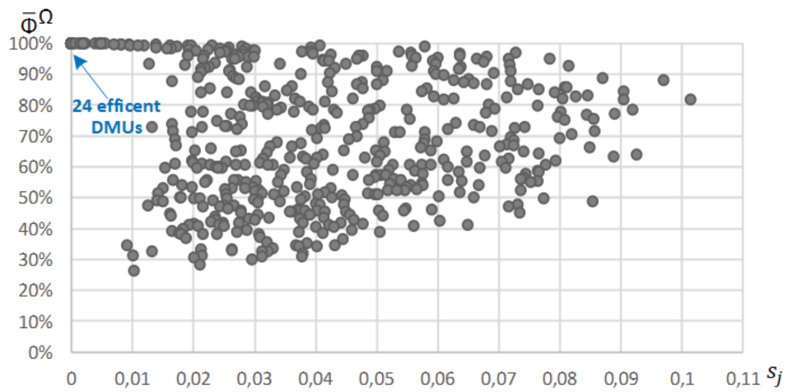

Figure 3: Mean and SD of efficiency of system configurations against different product mixtures.

Some system configurations show a high average efficiency indeed, but a relatively heavy fluctuation of the efficiency values as well. Thus these DMUs are less robust (e.g. $\bar{\Phi}^{316} \approx 91\%$, $s_{316} \approx 7\%$; $\bar{\Phi}^{89} \approx 89\%$, $s_{89} \approx 9\%$). The efficiency value of system configuration 91 ($\bar{\Phi}^{91} \approx 82\%$) reacts the most sensitively to variations of the product mixture with a standard deviation of $s_{91} \approx 10{,}2\%$. As already mentioned, 24 DMUs are efficient and also expose no deviance in efficiency against variation in the product mix (top left corner in Figure 3). These DMUs represent 960 (24 x 40) experiments, since every DMU/product mix combination has to be conducted as one single simulation run. Now knowing which experiments are efficient in terms of input and output criteria (used within the DEA model), we can start a visual analysis of the experiments regarding their relation to subsequent performance measures. Figure 4 shows a matrix of scatterplots showing selected simulation result parameters, in which efficient runs are highlighted in blue.

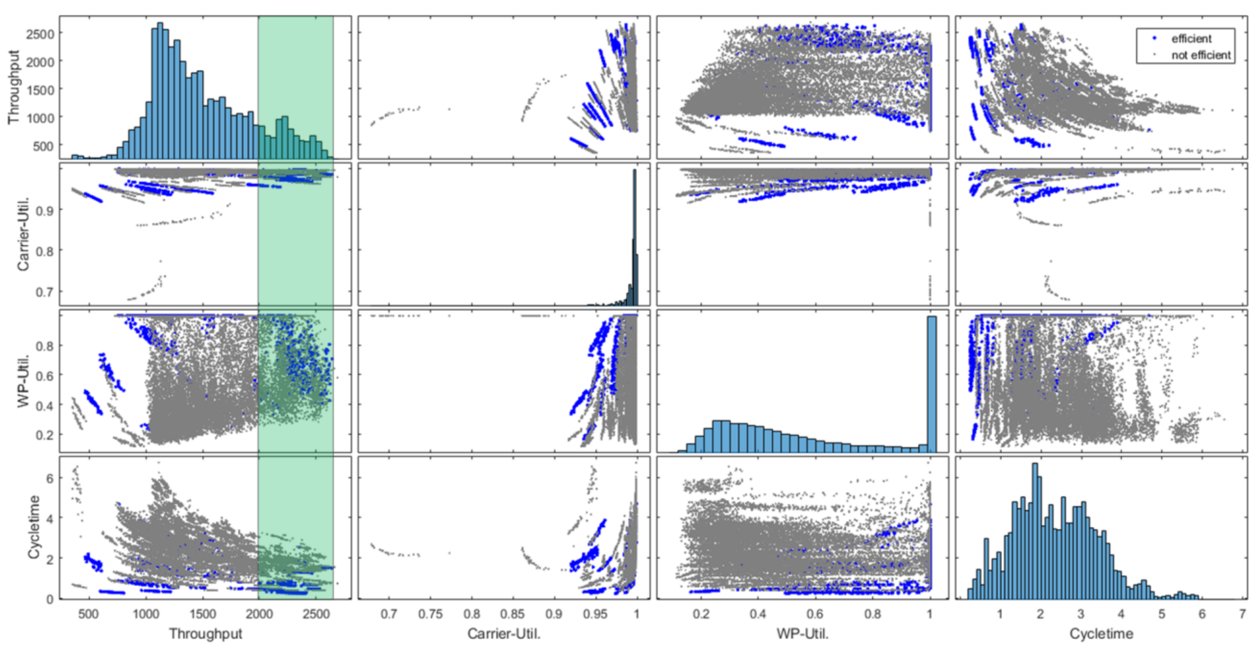

Figure 4: Matrix scatterplot of selected output performance measures.

Investigating corresponding factors to analyze what factor values lead to efficient runs over the complete dataset would not be feasible since the efficient runs are distributed over a large scale of system performance. Therefore, we first have to further reduce the area of investigation. If we take a look at the distribution of the efficient runs regarding the throughput, it is noticeable that there are different blocks of different levels of throughput. The first one is at ca. 500, the next between ca. 1000 – 1300, and the last one at ca. 2000 – 2600. The latter is one chosen for further analysis and is highlighted by a green strip. Efficient simulation runs in this subset also yield a low cycle time and a high utilization of carriers. Average workplace utilization for the most part is between 40% and 80%, so surprisingly for runs with high throughput, efficiency seems not being correlated with high utilization of the workplaces.

Because the NOLH design method that we used to create our experiments guarantees desirable features like orthogonality and balanced parameter values, we are able to analyze filtered subsets without any biased effects. To investigate corresponding input parameters of the subset highlighted in Figure 4 and the difference between efficient and not efficient experiments, we visually discriminated the subset into both categories, which is shown in Figure 5. Here, we selected all factors from the DEA and also included some other factors that are assumed to be most influential (which can be found through preceding analysis like for example visual correlation analysis). The so called radarplots in this figure mark the median and quartiles for each parameter. In between the quartiles are 50% of all observations (simulation experiments). So if quartiles lay close together, the corresponding parameter value is significant in the subset, which in turn represents good system performance in terms of our prior selected subset (Figure 4, highlighted green strip). More specifically, if a parameter value is equally significant in both left and right side of Figure 5, we can assume that this parameter value is most likely a necessary prerequisite to reach the output performance of the selected subset in the first place. On the other hand, if quartiles are very broad so that a factor value is rather equally distributed, the effects of this factor to the selected subset is presumed to be small. Note that both plots show subsets with good system performance but differ in efficiency. Therefore, we have to find differences in the left and right plot in order to conclude which system configurations or input parameter value combinations, respectively, lead to good and efficient system performance.

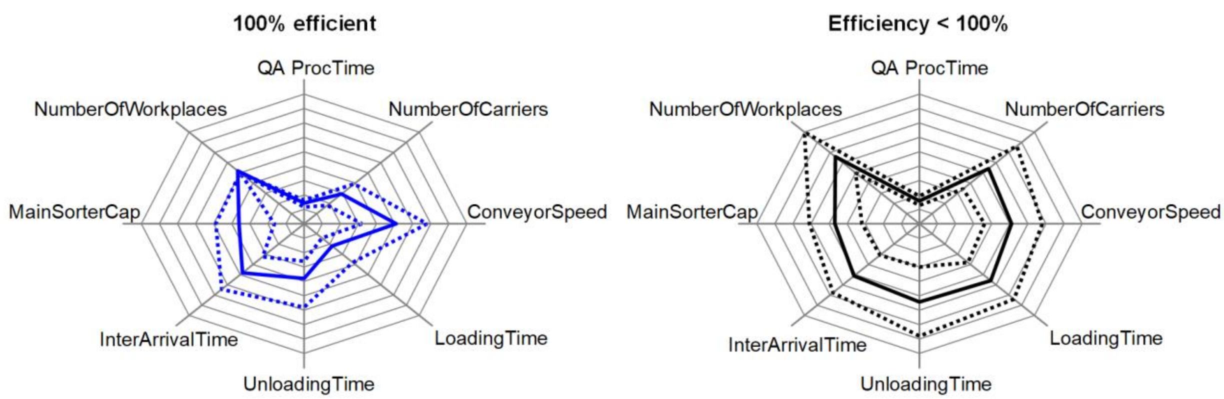

Figure 5: Radarplots of selected factors for efficient (left side) and not efficient (right side) experiments.

For example, both plots show a similar shape for the factor *QA_ProcTime*, which is the variance of the processing time for quality assurance station. Since the shape is the same for both plots and the parameter value is very significant, one might conclude that this factor is influential to the outcome of the system performance in the selected subset. For factors that have a significant shape but a very different for the left and right plot, we can conclude that they are important for efficiency. This is for example the case for factor *NumberofCarriers*, that is rather equally distributed for the non-efficient plot, but form a more distinct characteristic for efficient experiments.

Now that we have a first impression of how relevant input factors distribute, we want to further investigate the exact parameter values and possible factor value combinations that lead to efficient system behavior. For this purpose, we trained a binary decision tree on the data subset in order to build a model that can map the relation between input factors and efficiency in detail. The nodes of the tree represent a specific input factor value, the leafs or classes of the tree represent an outcome classification in terms of efficient or not efficient. Hence, each tree branch represents an if-then-rule that describes how to reach efficiency. A visualization of the tree is shown in Figure 6 (top frame). Note that the higher a split decision occurs in the tree, the more important the corresponding parameter is for the discrimination between efficient and non-efficient experiments. Therefore, *LoadingTime* can be leveraged in order to make the system performance (according to the dimensions of output parameters shown in Figure 4) efficient. Interestingly, *QA_ProcTime* does also have a big influence on whether or not an experiment is efficient, despite Figure 5 showed a similar shape for efficient and non-efficient runs for this factor and we therefore assumed that a low value for QA_ProcTime is most likely a necessary prerequisite to even reach the output performance of the selected subset. Looking at the tree, we can conclude that even values for *QA_ProcTime* above the threshold of 117s can lead to efficiency, if it occurs in combination with *InterArrivalTime* being <32s. Those interactions of factors, not all of which might be apparent from a comparison of radarplots alone, can be detected by tree visualization. Some of the factors in Figure 5 that appeared to be important are not present in the tree. This is for example the case for *NumberOfWorkplaces*. This occurs due to the fact that those factors are subsumed by other factors with which they are correlating in this specific data subset.

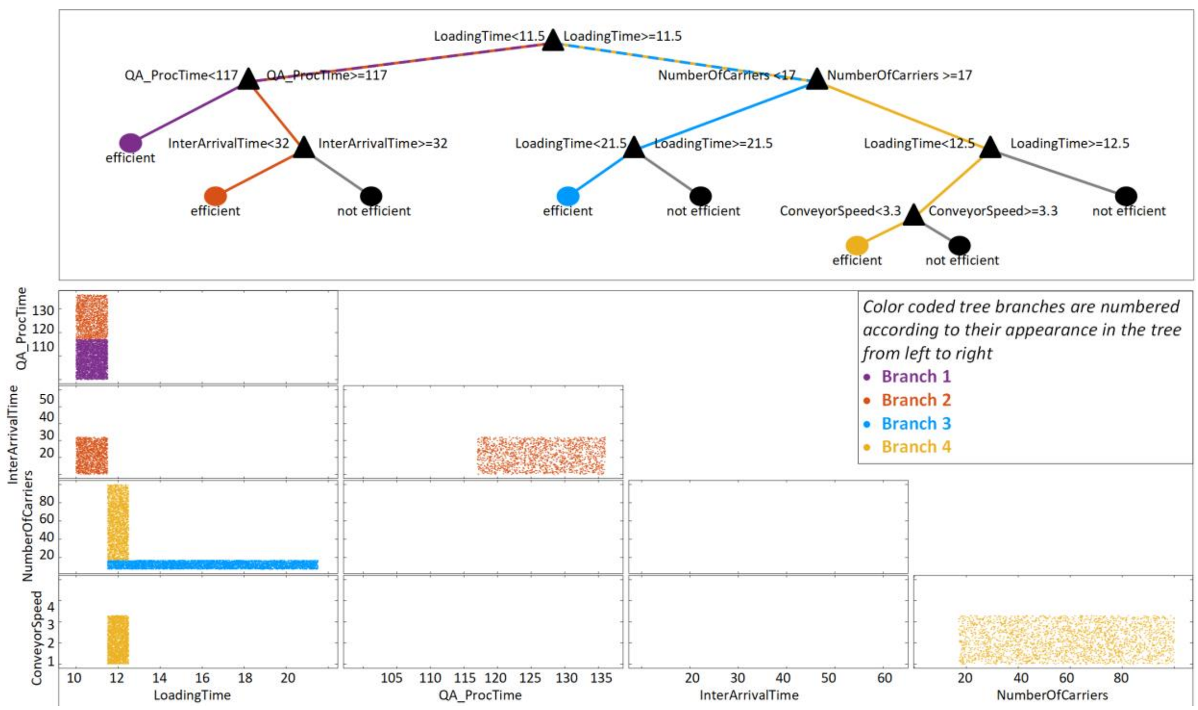

Figure 6: Decision tree (top frame) and decision boundaries visualization (bottom frame).

Furthermore, we color-coded the tree branches that lead to efficient classification and implemented a filter to corresponding if-then-rules and highlighted the related data in a scatter plot matrix (Figure 6 bottom frame). This method allows to visualize the decision boundaries of the tree, which represent the boundaries of factor values required to say within in order to reach efficiency. This enables a simpler exploration of factors and value boundaries that are affected by a tree branch, especially for factors that appear in the tree more than once, which is the case for *LoadingTime.* For example, the blue branch only affects two factors (*LoadingTime* and *NumberOfCarriers*), whereas the orange branch requires a combination of three factors (*LoadingTime*, *InterArrivalTime*, and *QA_ProcTime*). Subsequent analysis, for example on other prior selected subsets could be carried out to further investigate the models' behavior regarding the relation between performance and efficiency. We successfully derived knowledge about the system by means of interactive and visual analysis. Of course, derived hypothesizes can be additionally validated through additional experiments.

5 CONCLUSION, LIMITATIONS AND FUTURE WORK

In this paper, we demonstrated how DEA, a methodology for multi-criteria efficiency measurement, can be applied on large quantities of simulation output generated through data farming. We combined DEA with our existing visual analytics based knowledge discovery process for manufacturing simulation. For this purpose, artificially generated input and output data was exemplarily used for measuring the efficiency of a virtual manufacturing system. This allows for distinct efficiency evaluations of specific systems settings, which finally indicate most productive configurations. Since DEA is very rarely used in combination with simulation based applications, our approach brings an additional viewpoint for simulation data analysis in terms of investigating not only performance, e. g. in terms of cycle times and utilization, but also the productive efficiency of each experiment conducted. Furthermore, our approach might be more appealing to people who are not “simulation experts”, since an interactively designed and visually aided analysis process is more user-friendly. For future work, we want to apply this methodology on even more complex, real-world-models using empirical data. In that context it will be of major importance to discuss more intensively the different types of DEA models, particularly the choice between CCR- (constant returns to scale) and BCC models (variable returns to scale) (Banker et al. 1984) and hybrid returns to scale (Podinovski et al. 2014), respectively, as well as the incorporation of weight restrictions to refine value judgements between certain inputs and/or outputs (Roll et al. 1991). Furthermore, when using DEA to improve decision making in real life production systems, it is necessary to derive the input and output set more comprehensively from the decision makers’ objectives (Afsharian et al. 2013) and, thus, embed DEA within decision theory (Dyckhoff 2017).

REFERENCES

Afsharian, M., H. Ahn, and L. Neumann. 2013. "Generalized DEA: An Approach for Supporting Input/Output Factor Determination in DEA." *Benchmarking: An International Journal* 23:1892–1909.

Aminuddin, W. M. W. M., and W. R. Ismail. 2016. "Integrated Simulation and Data Envelopment Analysis Models in Emergency Department." In *The 2016 UKM FST Postgraduate Colloquium,* edited by K. Ibrahim et al., 1–7. Melville, New York: AIP Publishing.

Banker, R. D., A. Charnes, and W. W. Cooper. 1984. "Some Models For Estimating Technical And Scale Inefficiencies In Data Envelopment Analysis." *Management Science* 30:1078–1092.

Charnes, A., W. W. Cooper, and E. Rhodes. 1978. "Measuring the Efficiency of Decision Making Units." *European Journal of Operational Research* 2:429–444.

Cooper, W. W., L. M. Seiford, and K. Tone. 2006, *Introduction to Data Envelopment Analysis and Its Uses. With DEA-Solver Software and References,* 1st ed. Springer Science+Business Media Inc: Boston, MA.

Dyckhoff, H. 2017. "Multi-Criteria Production Theory: Foundation of non-financial and Sustainability Performance Evaluation." *Journal of Business Economics* 23:1892.

Dyson, R. G., R. Allen, A. S. Camanho, V. V. Podinovski, C. S. Sarrico, and E. A. Shale. 2001. "Pitfalls and Protocols in DEA." *European Journal of Operational Research* 132:245–259.

Elmegreen, B. G., S. M. Sanchez, and A. S. Szalay. 2014. "The Future of Computerized Decision Making." In *Proceedings of the 2014 Winter Simulation Conference,* edited by A. Tolk et al., 943–949. Piscataway, New Jersey: IEEE.

Fayyad, U. M., G. Piatetsky-Shapiro, and P. Smyth. 1996. "From Data Mining to Knowledge Discovery in Databases." *AI Magazine* 17:37–54.

Fazli, M. H., W.-L. Li, and T. Murata. 2012. "Vehicle Requirement Analysis of an AGV System using Discrete-Event Simulation and Data Envelopment Analysis." In *8th International Conference on*

Computing Technology and Information Management, edited by K. D. Kwack, 819–823. Piscataway, NJ: IEEE.

Feldkamp, N., S. Bergmann, and S. Strassburger. 2015a. "Knowledge Discovery in Manufacturing Simulations." In *Proceedings of the 3rd ACM SIGSIM Conference on Principles of Advanced Discrete Simulation,* edited by S. J. E. Taylor, N. Mustafee and Y.J. Son, 3–12. New York, NY, USA: ACM.

Feldkamp, N., S. Bergmann, and S. Strassburger. 2015b. "Visual Analytics of Manufacturing Simulation Data." In *Proceedings of the 2015 Winter Simulation Conference,* edited by L. Yilmaz et al., 779–790. Piscataway, New Jersey: IEEE.

Feldkamp, N., S. Bergmann, S. Strassburger, and T. Schulze. 2017. "Knowledge Discovery and Robustness Analysis in Manufacturing Simulations." In *Proceedings of the 2017 Winter Simulation Conference,* edited by W.K.V. Chan et al., 3952-3963, Piscataway, New Jersey: IEEE.

Horne, G. E., and T. E. Meyer. 2005. "Data Farming: Discovering Surprise." In *Proceedings of the 2005 Winter Simulation Conference,* edited by M. E. Kuhl et al., 1082–1087, Piscataway, New Jersey: IEEE.

Keim, D. A., F. Mansmann, J. Schneidewind, J. Thomas, and H. Ziegler. 2008. "Visual Analytics: Scope and Challenges." In *Visual Data Mining: Theory, Techniques and Tools for Visual Analytics,* edited by S. Simoff, M. H. Boehlen and A. Mazeika. 1st ed. Berlin, Heidelberg: Springer.

Kleijnen, J. P.C., S. M. Sanchez, T. W. Lucas, and T. M. Cioppa. 2005. "State-of-the-Art Review: A User's Guide to the Brave New World of Designing Simulation Experiments." *INFORMS Journal on Computing* 17:263–289.

Koopmans, T. C. 1951. "Analysis of Production as an Efficient Combination of Activities." In *Activity Analysis of Production and Allocation,* edited by T. C. Koopmans. New York: Wiley.

Mahfouz, A., and A. Arisha. 2015. "An Integrated Lean Assessment Framework for Tyre Distribution Industry." In *Proceedings of the 2015 Winter Simulation Conference,* edited by L. Yilmaz et al., 3196–3197 , Piscataway, New Jersey: IEEE.

McMullen, P. R., and G. V. Frazier. 1998. "Using Simulation and Data Envelopment Analysis to Compare Assembly Line Balancing Solutions." *Journal of Productivity Analysis* 11:149–168.

Min, H., and B.-I. Park. 2008. "Hybrid Data Envelopment Analysis and Simulation Methodology for Measuring Capacity Utilisation and Throughput Efficiency of Container Terminals." *International Journal of Logistics Systems and Management* 4:650–672.

Pareto, V. 1909, *Manuel d'économie politique*, Librairie Droz: Genève 12 (11, rue Firmin-Massot 1211).

Podinovski, V. V., I. Ismail, T. Bouzdine-Chameeva, and W. Zhang. 2014. "Combining the Assumptions of variable and constant Returns to Scale in the Efficiency Evaluation of secondary Schools." *European Journal of Operational Research* 239:504–513.

Roll, Y., W. D. Cook, and B. Golany. 1991. "Controlling Factor Weights in Data Envelopment Analysis." *IIE Transactions* 23:2–9.

Sanchez, S. M. 2000. "Robust Design: Seeking the Best of all Possible Worlds." In *Proceedings of the 2000 Winter Simulation Conference,* edited by J. A. Joinesm et al., 69–76, Piscataway, New Jersey: IEEE.

Sanchez, S. M. 2014. "Simulation Experiments: Better Data, Not Just Big Data." In *Proceedings of the 2014 Winter Simulation Conference,* edited by A. Tolk et al. , 805–816. Piscataway, New Jersey: IEEE.

Seiford, L. M. 1997. "A Bibliography for Data Envelopment Analysis (1978-1996)." *Annals of Operations Research* 73:393–438.

Thomas, J. J., and K. A. Cook. 2005, *Illuminating the Path. Research and Development Agenda for Visual Analytics,* 1st ed. IEEE Computer Society: Los Alamitos, California.

Vieira, H., S. M. Sanchez, K. H. Kienitz, and M. C. N. Belderrain. 2011. "Improved efficient, nearly orthogonal, nearly balanced mixed designs." In *Proceedings of the 2011 Winter Simulation Conference,* edited by S. Jain et al., 3600–3611. Piscataway, New Jersey: IEEE.

Weng, S. J., B. S. Tsai, L. M. Wang, and C. Y. Chang. 2011. "Using Simulation and Data Envelopment Analysis in Optimal Healthcare Efficiency Allocations." In *Proceedings of the 2011 Winter Simulation Conference,* edited by S. Jain et al., 1295–1305. Piscataway, New Jersey: IEEE.

AUTHOR BIOGRAPHIES

NICLAS FELDKAMP holds bachelor and master degrees in business information systems from the University of Cologne and the Ilmenau University of Technology, respectively. He is currently working as a doctoral student at the Department of Industrial Information Systems of the Ilmenau University of Technology. His research interests include data science, business analytics, and industrial simulations. His email address is niclas.feldkamp@tu-ilmenau.de.

SÖREN BERGMANN holds a Doctoral and Diploma degree in business information systems from the Ilmenau University of Technology. He is a member of the scientific staff at the Department for Industrial Information Systems. His research interests include generation of simulation models and automated validation of simulation models within the digital factory context. His email address is soeren.bergmann@tu-ilmenau.de.

ERIK BORSCH Erik Borsch holds a bachelor and master degree in industrial engineering for the Ilmenau University of Technology. He is currently working as doctoral student at the Department of Sustainable Production and Logistics Management at Ilmenau University of Technology. His research interests include efficiency analysis, especially Data Envelopment Analysis, and municipal solid waste management. His email address is erik.borsch@tu-ilmenau.de.

MAGNUS RICHTER is an Academic Assistant at the Department of Sustainable Production and Logistics Management, Ilmenau University of Technology. He holds a Diploma degree in Business Administration from RWTH Aachen University and a Doctoral degree in Economics and Social Sciences from Ilmenau University of Technology. His research is based on activity analysis, service science and operations management, focusing on waste disposal processes and customer interaction. His current research topics include DEA models for manufacturing systems. His email address is magnus.richter@tu-ilmenau.de.

STEFFEN STRASSBURGER is a professor at the Ilmenau University of Technology and head of the Department for Industrial Information Systems. Previously he was head of the "Virtual Development" department at the Fraunhofer Institute in Magdeburg, Germany and a researcher at the DaimlerChrysler Research Center in Ulm, Germany. He holds a Doctoral and a Diploma degree in Computer Science from the University of Magdeburg, Germany. His research interests include distributed simulation, automatic simulation model generation, and general interoperability topics within the digital factory context His email address is steffen.strassburger@tu-ilmenau.de.

RAINER SOUREN is a professor at the Ilmenau University of Technology and head of the Department of Sustainable Production and Logistics Management. Previously he was an Academic Assistant and Assistant Professor at RWTH Aachen University and Associate Professor at Bremen University. He holds a Doctoral and Diploma Degree in Business Administration from RWTH Aachen University. His research interests include production theory, performance measurement and sustainable production and logistics systems. His email address is rainer.souren@tu-ilmenau.de.