International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:11, No:1, 2017

Procedure Model for Data-Driven Decision Support Regarding the Integration of Renewable Energies into Industrial Energy Management

M. Graus, K. Westhoff, X. Xu

***Abstract*—**The climate change causes a change in all aspects of society. While the expansion of renewable energies proceeds, industry could not be convinced based on general studies about the potential of demand side management to reinforce smart grid considerations in their operational business. In this article, a procedure model for a case-specific data-driven decision support for industrial energy management based on a holistic data analytics approach is presented. The model is executed on the example of the strategic decision problem, to integrate the aspect of renewable energies into industrial energy management. This question is induced due to considerations of changing the electricity contract model from a standard rate to volatile energy prices corresponding to the energy spot market which is increasingly more affected by renewable energies. The procedure model corresponds to a data analytics process consisting on a data model, analysis, simulation and optimization step. This procedure will help to quantify the potentials of sustainable production concepts based on the data from a factory. The model is validated with data from a printer in analogy to a simple production machine. The overall goal is to establish smart grid principles for industry via the transformation from knowledge-driven to data-driven decisions within manufacturing companies.

***Keywords*—**Data analytics, green production, industrial energy management, optimization, renewable energies, simulation.

I. Modern Industrial Energy Management

THE great approval of the participants of the United Nations Climate Change Conference in 2015 seemed to indicate that the main part of the world is aware of the necessity to change the way that we handle energy. There have already been lots of efforts to reduce greenhouse gas emissions in several countries. One example is the energy turnaround in Germany, which is one large project to break well-known structures and patterns of a whole society on a scale of only a few decades. However, there is still a need to establish more extensive concepts for green industrial technologies and sustainable production to overcome the challenges of a new climate policy. [1] It is necessary to use synergies of different streams to tackle these biggest challenges of our society. The role of industry in the sustainability movement is specifically interesting. While renewable energies are on the rise and cities are reworked/modernized with the goal of energy efficiency, the inclusion of the industry in the so-called future smart grid seems to be a major open task. The idea of a demand side management is quite old, but there are only a few successful examples put into practice so far. However, industry is not only influenced by the sustainability movement. A further central influence on the industry is the development towards the internet of things and the fourth industrial revolution or how it is called in Germany: Industrie 4.0. There is a natural synergy of the energy turnaround and Industrie 4.0 within the area of industrial energy management. [2] The energy management tends to include more and more data enabling the establishment of new concepts in production management. In this article, a procedure for data analytics to include the aspect of renewable energies into the industrial energy management will be introduced.

In general, there are three dimensions of management: normative, strategic, and operational management. [3]

In energy management, the operational dimension is very interesting and complex since it is closely correlated to the production management itself. Currently operational energy management in industry ends mostly with a monitoring of the energy consumption and possible measures like invest into more energy efficient machines and buildings, which are more strategic decision. The integration of energy efficiency in operational processes is a much more complicated topic that is mainly been avoided. However, there are a number of scientific works about demand-side management and analytics of energy consumption data in manufacturing. The integration of energy consumption costs into the production scheduling on the example of a single machine was focused in [4], while [5] shows a Key Performance Indicators (KPI) based approach. One weakness of common concepts is that they are mainly knowledge-driven in the sense that an expert builds a model that solves the problem in a theoretical environment. One characteristic example is the neglect of the gap between production planning and the actual production process. In contrast to a knowledge-driven approach one can develop data-driven concepts, which become increasingly popular in many parts of the economy due to the digitalization [6]. As a

M. Graus is with the Information Management Department, FIR at RWTH Aachen University, 52074 Aachen, Germany (phone: +49241 47705-514; e-mail: Marcel.Graus@fir.rwth-aachen.de).

K. Westhoff is with the Technical Communication and Computer Science Department, RWTH University, 52056 Aachen, Germany (e-mail: katalin.westhoff@rwth-aachen.de).

X. Xu is with the Industrial Engineering Department, Zhejiang University, Hangzhou, Zhejiang, 310027, China (e-mail: xuxin1991@zju.edu.cn).

consequence, for the modelling, the general conditions are determined a priori by an expert, and the main properties of the model are deduced from the empirical data without further influence from any expert. Moreover, most available works do not focus on a complete adaptive process that performs the transformation from a data basis to information used for a decision support. [7] In contrast, the focus in this article is on the process itself and resign a detailed explanation of the techniques used for each and every step. In order to simplify the individual steps, a consideration of only one simple machine is preferred, although the introduced concept can be applied for more complex machines and full processes as well. The application of the process for strategic energy management decisions in industry should help to estimate and release the potential of sustainable production concepts. The first barrier on this way is the need to convince the decision-maker in strategic management decisions based on facts that are specific for the regarded company. The example of the integration of fluctuating energy prices in the production planning and control is not well understood so far, since today's investigations are mostly hypothetical studies, while a case specific study is needed to convince a decision-maker in an enterprise. Thus, a suitable procedure to realize the idea of case-specific analysis of the potential becomes necessary.

II. Approach

Data Analytics is a process of transformation from data to information that is useful for decision making. [8] This process is specific for every application domain and every use case. In the following, an approach for this process is regarded in the domain of the industrial energy management as it is introduced in [8], [9] shown in Fig. 1. It starts with the building of a data model which abstracts the data basis and allows further analytics. The analysis step is the key to understand to data basis and deduce insights from the empirical data. There, one needs to distinguish between the analysis of historical data and real time analysis [9]. In the simulation step, the line of sight is reversed to realize predictions that respect current trends and possible active intervention in the process in the factory. Based on the analysis and simulation in the step of optimization, the rule of intervention as well as the best of all scenarios is determined. This approach is applied to the problem of a strategic decision of integrating volatile energy prices within the planning of a production machine. This could be initiated by a change of the electricity contract from a standard tariff to volatile prices that are geared on the generation of renewable energy. The concept is exemplified on a simple scenario. A printer, as one can find one in every office, was identified as an interesting machine which has several similarities to an ordinary production machine but is a lot easier to understand without well-founded knowledge of a certain industry. There are in general machine data which indicate the current state of a printer but is not always available. After installing a sensor, information is gathered about the energy consumption like the power per second. For this use case, the data model is not crucial, since the data basis is sufficiently simple. For other industrial machines and processes, the complex data basis in the enterprise requires intensive thoughts about the data model. Furthermore, optimization of the performance of the analytic steps is not required, since a strategic and non-operational decision problem is tackled, which would need results immediately. This would induce a need of a suitable IT infrastructure to perform the calculations. [8] As data basis, a sample of power per second from a printer in an office and appropriate 15 minutes and spot market energy prices from EPEX gathered in three months are used. Since the printer itself is not energy intensive and not working to capacity, the potential is limited and the specific values of power consumptions are not important. Thus, the concrete data values are neglected in the following and focused instead on the concept itself.

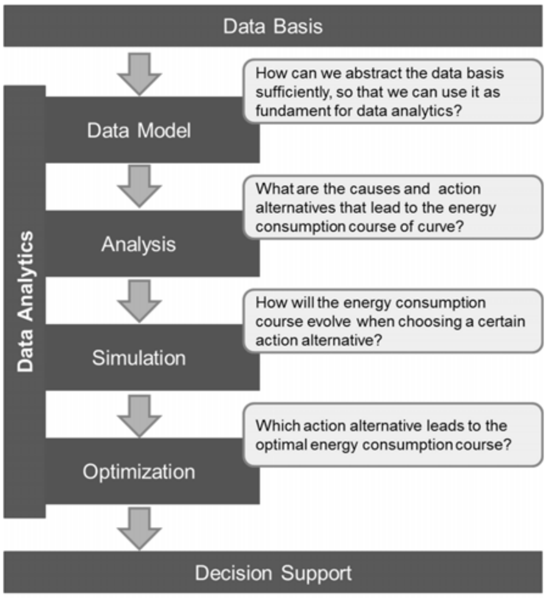

Fig. 1 Data Analytics Approach in [8], [9]

III. Analysis

The analysis step is very important and will need the most effort in many applications. The goal of the analysis is to understand the energy consumption course and learn from historical data causes and effects, which is the basis for arriving good predictions in the simulation step. If there is any description of the phenomena via equations like usual in applied mathematics in physics, the analysis step is the one, where one creates the model and investigate its properties. However, in manufacturing companies many phenomena are too complex and uncertain to describe them a priori with some formula. Thus, it is useful to start with a more dynamic approach where a part of the model structure will be learned directly from the data. These models can be arbitrarily complex in analysis and implementation. In order to focus on the procedure itself and the fact that just an estimate of a certain quality is wanted, the analysis is solved with a simple two-step approach. In the first step, the time series is segmented into a number of intervals through a change point detection algorithm. In the second step, the segments are classified via a k-means algorithm with the mean values of the power within every segment. Fig. 2 shows the segmentation of a classification into two classes (black and white).

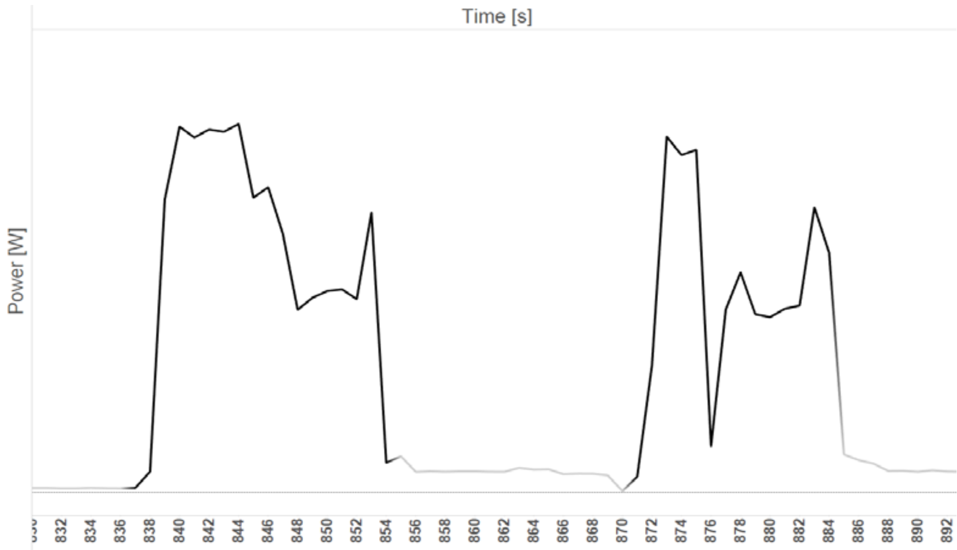

Fig. 2 Time Series Analysis

Since there is later in the simulation and optimization a distinction only between the kind of job and work approximately with a finite number of energy consumption level, this approach leads to sufficient results. In addition, this approach has the advantage that it is also applicable for data streams. One possibility in the implementation is the usage of Apache Spark Streaming which supports with the concept of microbatches this two-step approach. Further, the notion of similarity in the second step could be improved to achieve deeper insights into the energy patterns inside every machine state. One approach that was applied successfully for the similarity model can be found in [10]. Typical challenges that one has to tackle are the definition of a meaningful metric in the space of the segments since they are in general not of the same length and handling offset translation, amplitude scaling, and outliers. There are a lot of other approaches especially from machine learning to realize the time series segmentation. Another aspect that needs to be focused in other use cases is the connection to other data sources like machine and planning data.

IV. Simulation

In industry, prediction is always a difficult task, since in most of the companies the processes are not completely digitalized and automatized, and the factor human adds even more uncertainty. This induce that a number of methods for predictions are not suitable in this context. One example of a failing strategy is an ordinary regression approach to predict energy consumption on a long run. Instead the notion of simulation seems to be a lot more appropriate, since simulation can be based on stochastic that models the uncertainty which is useful for the sake of prediction in a wide range of application domains. The main idea here is that one first needs to find a model which explains sufficiently the effects of the past. Then, the model can be used to forecast the prospective course, if the conditions stay unchanged. In addition, it is possible to adjust the model and create what-if-scenarios as a result.

A. Parameter Estimation

For this example, an event based approach has been chosen that builds upon the analysis step which provides information about energy events from the empirical data. An energy event e is a tuple (x, d, E), where x corresponds to the starting time of the event, d is the event duration, and E is the energy consumption level. While there are only a finite number of energy consumption level as subset of the positive real numbers in the model, the starting time and duration have an infinite but countable state space dimension with the natural numbers, since time has been discretized. The assumption is, that there is no dependency between the components of a single event tuple.

In the model, d is a time independent random variable and x is determined by $x_{ts+1} = x_{ts} + d_{ts} + 1$ where $ts = 1, \dots N$ indicates the time step from one event to another. The energy consumption level E is a time depended random variable. Therefore, the first task is to determine the distribution for d (Fig. 3) and its parameter of the model based on the results of the analysis step. For the parameter estimation, maximum likelihood approaches are used.

One weakness of a temporary measurement is that the potential of integrating energy demand into the scheduling process can only be calculated on the basis of the limited sample of empirical data, and thus, only a small fraction of the phenomena could be observed. Therefore, for a statement about longer time intervals or other conditions than in the example that prevail during the measurement a scale-up of the data is required. In the case of the printer, one could regard the scenario where the printer is used closed to capacity. It is important that this mapping preserves the essential properties of the empirical data.

B. Monte Carlo Simulation

Once the model is determined from the measurement, one needs to predict the behaviour of the system respectively the machine, when there is no active intervention in the process. Since a stochastic approach was chosen for the analysis and simulation step one gets in prediction in the form of a curve that will in a sense most likely happen and a confidence band indicating in which area the future observations will appear with high certainty (Fig. 4). If the modelling and analysis was successful, one gets closely exact results with increasing number of simulations by the law of large numbers.

Based on the simulation, one can calculate some KPIs to prepare a management decision. If the energy prices are plugged in, the energy costs can be deduced on the basis of the simulation of the energy consumption. Moreover, one obtains a mean waiting time. The next step is to predict the observations if one intervenes actively in the process based on the certain rule. This rule is an optimization itself which is presented next.

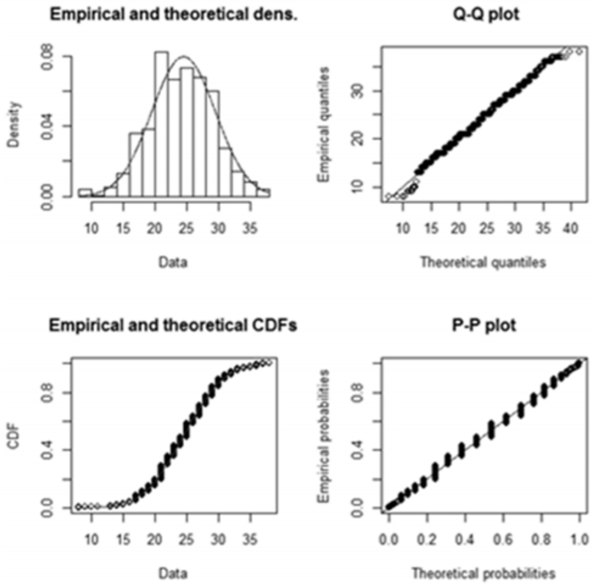

Fig. 3 Distribution Fitting

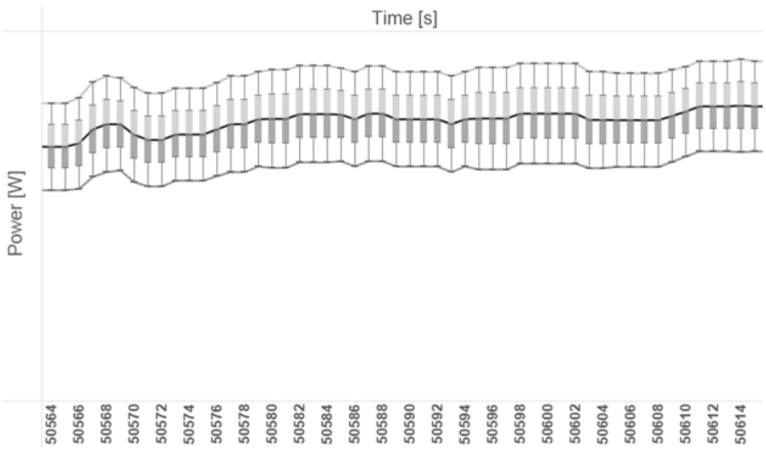

Fig. 4 Simulation of Energy Consumption

V. Optimization

The optimization model is used to find the optimal scheduling for the machine in the case that the current energy price is included. Regarding the optimization, the notion of energy events leads to false association, since a decision-maker needs to decide now in the sense of scheduling when an event starts. Thus, the energy events become jobs in this chapter. The task of scheduling as essential in production planning and includes always the challenge of balancing several targets in production management. Examples of production specific targets are adherence of delivery, minimizing production costs and throughput time. In the regarded example of the printer, the waiting time for the completion of a job corresponds to the throughput time and adherence of delivery, while the energy costs correspond to the production costs. The prioritization of the different targets is specific for every industrial domain and every company. In this article, a solution based on dynamic programming principles is presented; however, there are many other approaches to handle similar problems. A stochastic model can be found in [11], while heuristic is often used for the solution idea.

A. Modelling

The inputs for the optimization model are the energy price P_i per time unit, a job list with requesting time r_j, duration d_j and the power consumption E_j per time unit with respect to the different machine states. Another important parameter is the measure to trade production costs against waiting time. A time interval $[0, K_f]$ is regarded, which one can segment into time units of unit length $1, \dots, i, \dots, K$. Every job j is characterized by its requesting time r_j and the duration d_j. Further, every job has a certain power consumption level E_j. A variable has to be introduced to hold the scheduling of the jobs:

$$a_{ij} = \begin{cases} 1, & \text{if job } j \text{ is processed in time unit } i, \\ 0, & \text{else.} \end{cases} \quad (1)$$

Thus, the total energy costs can be expressed as

$$\sum_i \sum_j P_i E_j a_{ij}. \quad (2)$$

The waiting time for one job is

$$\min_i \{i: a_{ij} = 1\} - r_j. \quad (3)$$

This leads to the following mathematical problem:

$$\min_{a_{ij}} \left\{ \sum_i \sum_j P_i E_j a_{ij} + \theta \times \sum_j \left[\min_i \{i: a_{ij} = 1\} - r_j \right] \right\} \quad (4)$$

subject to:

$$\sum_j a_{ij} \leq 1, \quad (5)$$

$$\min_i \{i: \mathrm{a}_{ij} = 1\} \geq \mathrm{r}_j, \quad (6)$$

$$\max_i \{i: \mathrm{a}_{ij} = 1\} - \min_i \{i: \mathrm{a}_{ij} = 1\} = d_j, \quad (7)$$

$$a_{ij} \in \{0,1\} \text{ and } \sum_i a_{ij} = d_j, \quad (8)$$

$$\min_i \{i: \mathrm{a}_{ij} = 1\} < \min_i \{i: \mathrm{a}_{ij'} = 1\}, j < j'. \quad (9)$$

This formulation of the problem is kind of cumbersome if one tries to tackle it in this form.

B. Analysis and Implementation

One challenge is to handle the min/max terms in (4)-(9). A further analysis of the problem indicates that one has a formulation that allows smaller subproblems with the same structure. In order to solve this kind of optimization problem, it is obvious to apply a dynamic programming method. Due to this idea, the state-transition function can be defined

$$f(i,j) = min \begin{cases} f(i+1,j) + p_i E_0\,, \\ f(i+d_j, j+1) + \sum_{m=i}^{i+d_j-1} P_m E_j + \theta * (i - r_j). \end{cases} \quad (10)$$

The first term in (10) corresponds to the stand-by state, while the second term indicates that job j should be started. Calculating this function iteratively the optimal scheduling can be determined via reconstructing backwards in time as usual in dynamic programming. One example of the result is shown in Fig. 5. The upper graph shows the original scheduling, where white means that the machine is in stand-by and black means it is operating. The lower graph shows the scheduling after optimization, where most of the jobs are done in time intervals where the energy price is lower.

It is clear that the degree of change between both graphs depends on the factor θ and the fluctuation of the energy prices. In the end, the management has to specify the factor based on the company's prioritization. The decision to include volatile energy prices via a possible contract change can now be done based on KPIs like additional mean waiting time and saved energy costs. The main difference in the decision-making process is the fact that the decision can now be based on data and the actual processes in the factory. This approach behaves diametrical to a consideration where this accords with a transformation from a knowledge-driven to a data-driven decision process.

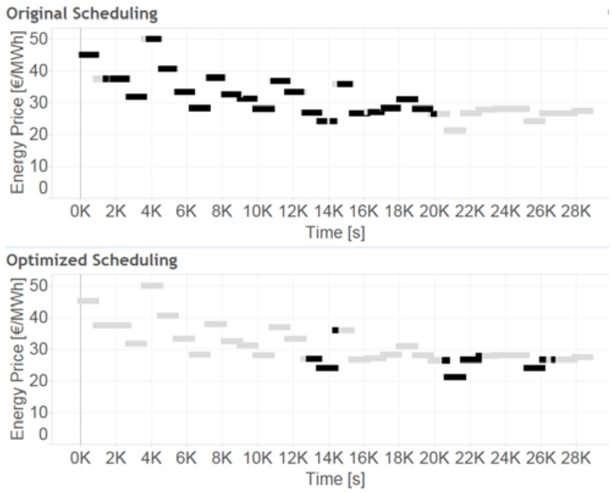

Fig. 5 Original and optimized scheduling

VI. Conclusion

A data analytics approach has been presented, that could be used for a decision support regarding integration of renewable energies into industrial energy management. The view of data analytics as a process leads to a division into four steps to lead from the data basis to a decision support. The process begins with data modelling as the abstraction of the data basis for the sake of analytics. The analysis step was used to understand the behaviour of the system from historical data before one can deduce predictions for the development with and without active intervention in the process. The final step is an optimization that realizes a rule of process intervention. In the end, the decision can be supported with KPIs based on the data analytics. This approach is executed on the example of a printer and the decision problem to change the electricity contract from a standard rate to a volatile energy price according to the energy spot market. The mathematics follows a stochastic approach for the analysis and simulation. The implementation was realized via R, MATLAB, and Tableau. As a result, the potential of coupling the printer to fluctuating energy prices must be rated as low. The reduced energy prices are the main reason today. Since the focus of the paper was to introduce the data analytics process for decision processes there are many possibilities to extend and improve the single steps of the process. For the transfer from the example of the printer to a production machine or process in factory, it will be necessary to start with a more complex data model that could be based for example on an industrial protocol like OPC-UA. Further, the analysis will become more complex, which can cause the necessity to introduce nontrivial machine learning approaches to learn the systems behaviour from the data. In an ordinary production, the optimization will be affected by a multidimensional balancing problem including a higher number of targets. The transfer to other decision problems that are not from strategic but operational nature like load management in production control higher requirements for the IT infrastructure and the introduction of parallelization ideas become necessary. Future research will focus on the transfer of the concept to the production management and operational decision problems.

ACKNOWLEDGMENT

This article arose during the work of M. Graus, within the framework of the research project FlAixEnergy (project number: 0325819A) funded by the Federal Ministry for Economic Affairs and Energy in Germany. The authors want to thank all donors, supporter and critics.

REFERENCES

[1] Bundesministerium für Wirtschaft und Energie (BMWi), „Energie der Zukunft – Vierter Monitoring-Bericht zur Energiewende – Langfassung“, online available at: https://www.bmwi.de/DE/Mediathek/publikationen,did=739122.html, 2015, p. 4 et seq.
[2] Acatech, Forschungsunion, “Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0”, online available at: https://www.bmbf.de/files/Umsetzungsempfehlungen_Industrie4_0.pdf, 2013, pp. 23-31.
[3] K. Bleicher, “Das Konzept Integriertes Management”, Frankfurt/Main, Campus Verlag Frankfurt/New York, 2004, pp. 80-84.
[4] Fadi Shrouf, Joaquin Ordieres-Meré, Alvaro García-Sánchez, Miguel Ortega-Mier, “Optimizing the production scheduling of a single machine to minimize total energy consumption costs”, in Journal of Cleaner Production 67, 2014, pp. 197-207.
[5] S. Bougain, D. Gerhard, C. Nigischer, S. Ugurlu, “Towards energy management in production planning software based on energy consumption as a planning resource”, in Procedia CIRP 26, 2015, pp. 139-144
[6] M. J. Schniederjans, D. Schniederjans, C. M. Starkey, „Business Analytics Principles, Concepts, and Applications: What, Why and How.”, Upper Saddle River, New Jersey, Pearson Education, 2014, pp. 96-99.
[7] N. Lin, „Applied Business Analytics: Integrating Business Process, Big Data, and Advanced Analytics”, Upper Saddle River, New Jersey, 2015, pp. XVI et seq.
[8] M. Graus “Integration einer Datenanalytik in Energieinformationssysteme produzierender Unternehmen”, in Smart Energy 2016, pp. 44-55.
[9] M. Roscher, M. Graus, “Energiewende in der Industrie”, in VDE Kongress 2016, Internet der Dinge, 07. -08.11.2016 in Mannheim, 2016.
[10] R. Agrawal, K. Lin, H. S. Sawhney, K. Shim, “Fast similarity Search in the Presence of Noise, Scaling, and Translation in Time-Series Databases”, in Proceedings of 21th International Conference on Very Large Data Bases. Zurich, 1995, pp. 490-500.
[11] N. Löhndorf, M. Riel, S. Minner, “Simulation Optimization for the Stochastic Economic Lot Scheduling Problem with Sequence-Dependent Setup Times”, in International Journal of Production Economics, Volume 157, 2014, pp. 170-176.

Marcel Graus studied mathematics at the University of Bonn and University Pierre and Marie Curie in Paris. Today, he works at the FIR at the RWTH Aachen University in the Information Management Department as scientist and project manager. His research focus is on the application of data analytics in the field of industrial energy management. He is currently working on his doctoral thesis in mechanical engineering at RWTH Aachen University about the topic of data analytics for energy information systems of manufacturing companies.

Katalin Westhoff studied technical communication and computer science and graduates with Master’s degree at the RWTH Aachen University. She worked as student assistant at FIR at the RWTH Aachen University with a focus on predictive analytics and the development of simulation models.

Xin Xu studies Industrial Engineering at Zhejiang University and is now working on his master’s thesis which he already started during his semester abroad at the RWTH Aachen University. In this period, he worked together with Mr. Graus in the area of optimization regarding the integration of fluctuating energy prices into the production planning. Before, he already completed his studies of mathematics with the bachelor’s degree at Zhejiang University.