Data Mining Techniques for Energy Efficiency Analysis of Discrete Production Lines

Rafal Cupek[1], Jakub Duda[1(✉)], Dariusz Zonenberg[2], Łukasz Chłopaś[2], Grzegorz Dziędziel[2], and Marek Drewniak[2]

[1] Institute of Informatics, Silesian University of Technology, Gliwice, Poland
{Rafal.Cupek, Jakub.Duda}@polsl.pl
[2] AIUT Sp. z o.o. (Ltd.), Gliwice, Poland
{dariusz.zonenberg, lukasz.chlopas, grzegorz.dziedziel, marek.drewniak}@aiut.com.pl

Abstract. Machine-level energy efficiency assessment supports the rapid detection of many technological problems related to a production cycle. The fast growth of data mining techniques has opened new possibilities that permit large amounts of gathered energy consumption data to be processed and analyzed automatically. However, the data that are available from control systems are not usually ready for such an analysis and require complex preparation – cleaning, integration, selection and transformation. This paper proposes a methodology for energy consumption data analysis that is based on a knowledge discovery application. The input information includes observations of the production system behavior and related energy consumption data. The proposed approach is illustrated on the use case of an energy consumption analysis that ws prepared for an automatic production line used in electronic manufacturing.

Keywords: Energy efficiency · Data mining · Manufacturing Execution System (MES) · Industrial communication · OPC UA (IEC 62541)

1 Introduction

As discrete production lines become more and more complicated, predictive maintenance activities have become a vital task for the engineers responsible for production support. Many potential technological and technical problems can be detected based on early symptoms that are noticeable in changes in energy consumption first [1]. Progressive faults can be caused by the aging or deterioration of the operating environment. Often progressive faults are noticed too late. Slow changes can especially be seen in greater energy consumption. So use of the timely monitoring of conditions and modern methods for fault diagnosis are required [2]. These provide the opportunity to detect maintenance or technological problems on the machine level by observing and processing information about energy consumption [3]. The current data can be compared with the information related to the energy consumption profiles in an appropriate production context. This comparison allows maintenance activities to be planned in advance and, as a consequence, reduces the losses related to production breakdowns.

N.T. Nguyen et al. (Eds.): ICCCI 2017, Part II, LNAI 10449, pp. 292–301, 2017.
DOI: 10.1007/978-3-319-67077-5_28

One of the main challenges related to predictive maintenance activities is the limited hardware resources that are available in PLC (Programmable Logic Controller) based control systems. Despite emerging hardware solutions that permit advanced data processing directly on the control system level [4], in most cases, the information related to energy consumption has to be sent to Business Intelligence systems for further processing. In such a case many problems related to the collection of real-time process information in distributed and large-scale data acquisition systems [5] have to be solved. The next issue is the discrimination of different production variants, which can be solved using a data mining approach [6].

This paper presents a new algorithm to automatically classify the energy consumption data that is relevant to machine production cycles. The proposed method is based on observing the behavior of a machine (control signals) and can be used in the case of mass-customized production. A predictive maintenance diagnostic system uses the results of the proposed clustering algorithm to automatically assess energy efficiency on the production station level. The results are further used for the early detection of machine faults that are visible by anomalies and changes in energy consumption. In the experimental part, the authors illustrate the proposed approach through an example of the automatic processing of information relevant to the production system behavior and its energy consumption that was created for a fully automatic electronic production line called "OKO" which is owned by the AIUT company.

The rest of this paper is organized as follows: the second section is focused on the data mining methods that can be applied for energy consumption information analysis with a special focus on multivariate discrete production systems. The third section presents the main components that are used in the proposed approach. Section four gives some details for the considered use case, and briefly describes the system that was prepared for the production line "OKO". The conclusions are presented in section five.

2 Application of Data Mining Methods for Energy Consumption Analysis

Han, Kimber et al. provide a definition of data mining as "as process of discovering interesting patterns and knowledge from large patterns of data" [7], also known as knowledge discovery from data, which they described as an iterative sequence of: data cleaning, data integration, data selection, data transformation, data mining, pattern evaluation, and knowledge presentation This definition directly applies to the research presented in this paper, and to the challenges of energy efficiency analysis in general. The presented research emphasizes data transformation and mining - building tools for efficient knowledge discovery.

The existing literature on data mining and energy efficiency often concentrates on electrical energy consumption in the context of buildings as presented by Zhao et al. in [8]. A detailed overview of the measures, methods and techniques for energy efficiency optimization in discrete manufacturing was provided by Duflou et al. in [8], while Cannata et al. presented and discussed an energy efficiency driven process in [9].

The initial goals for the presented research were to gather and analyze energy related measures from a real production line, by looking for anomalies and interesting patterns in energy consumption and efficiency. For this purpose, the "OKO" production line – a robotic line used in testing, total quality management (TQM) and final assembly of natural gas use GPRS telemetry devices was selected.

However, the authors wanted to abstract the analysis from this particular line - the findings, conclusions and developed tools should be universally applicable to discrete production lines, especially in the automotive industry. Another aim was to look for easily identifiable markers for abnormal energy consumption that can be determined on-the-fly in a production environment, using only PLC processing.

For the purpose of the research, the following algorithm (presented in Fig. 1) was proposed:

1. Gather production process data and energy efficiency data snapshots
2. Store the data snapshots in the source database
3. Cleanup the data and transform it from the source format (snapshots) to energy efficiency data
4. Apply data mining analysis
5. Provide the results to the clients

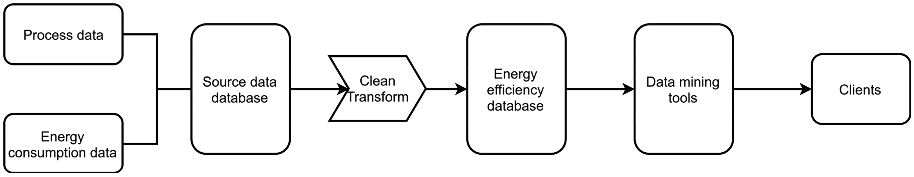

Fig. 1. Proposed algorithm for energy efficiency data analysis

To accomplish this, the communication setup presented in Fig. 2 was developed. The production data that was relevant to energy consumption and the actual measures (electrical energy ad compressed air) were read from PLC using an industrial network by an OPC UA compliant server. These values were made available in an Ethernet network using the OPC UA data access protocol, and read by client software with historian capabilities, which stored all of the relevant data in an SQL database. For efficiency and ease of use the Unified Automation OPC UA SDK and MS SQL database were used for the historian software.

Finally the gathered historical data can be batch processed in data mining environment – for this purpose RapidMiner was selected.

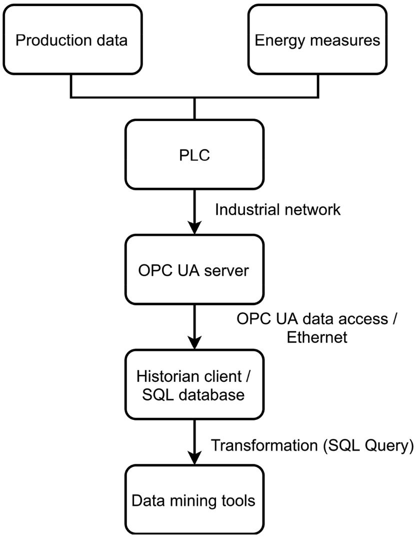

Fig. 2. Communication setup

Although this setup may look excessive at first (the OPC UA server could theoretically be integrated with the SQL historian for this particular line), it allows the data gathering process to be easily scaled by adding more PLCs (it is not unusual to have a hundred PLCs per production line) and distributed data servers while maintaining the central data processing point. As there are multiple telemetry devices processed in parallel in the OKO line at each moment (up to five in the testing stations, one in the laser station and up to 20 in the plotter), it is virtually impossible to determine energy consumption per product. Therefore, the decision was made to analyze the production cycles - especially the robot movements - instead of the product cycles. This approach supports comparison of the energy footprint for repeated actions over long periods of time, which should allow any anomalies or efficiency deterioration to be easily detected.

3 Energy Consumption Analysis for the Fully Automatic Electronic Production Line "OKO"

The data provided by the PLC control routines is just a snapshot of the state of the line devices and meters at a given point in time (as presented in Fig. 3). For electrical energy and compressed air it is the instantaneous consumption and the state of the global meters. Because these are not directly applicable for energy efficiency analysis, the data needs to be transformed into information that describes the production cycles (presented in Table 1).

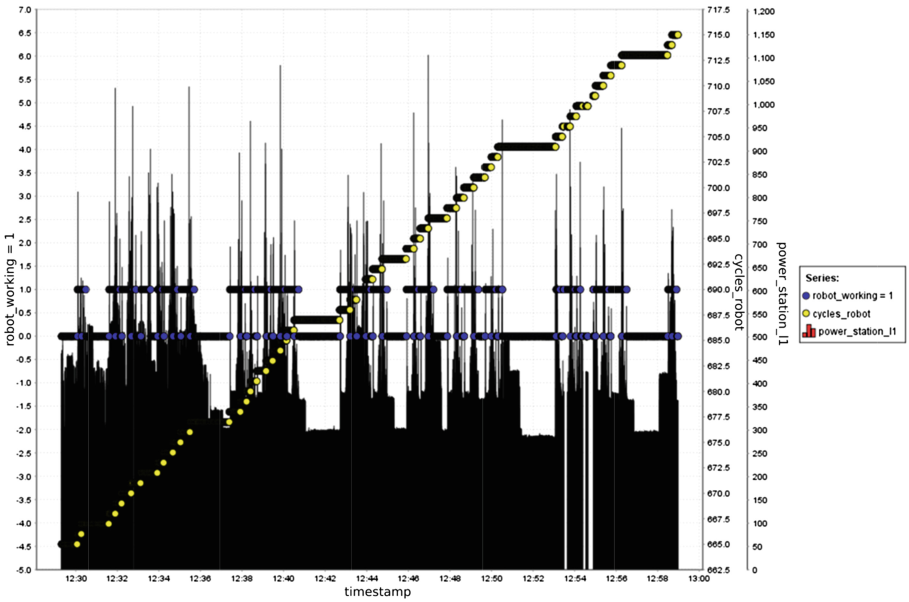

Fig. 3. Source data – instantaneous energy consumption, robot busy status and robot cycle counter shown as a function of time.

Table 1. Comparison of the available source data and the data required for a production cycle energy analysis.

Source data	Transformed data
Energy meter	Energy use per cycle
Air flow meter	Air use per cycle
Robot cycle counter	Cycle time
Device state bits	Robot movement type

4 Data Analysis by RapidMiner Tool

There are several ways to apply the data transformations that are required for analysis. One approach could be based on OPC UA events handling directly in historian software. However besides problems with data verification (source data would not be stored) this method also requires the production events to be coded into the client software as subscription event conditions, thus making it prone to errors and difficult to change, especially when a process is evolving.

Another possibility is to use the data mining environment logic conditions – RapidMiner provides many functional blocks for manipulating data that make it possible to perform all of the transformations this way. The main drawback for this approach is that the transformation method is tied to this particular environment.

Eventually SQL query was selected as the method for data transformation. This is a proven industry standard that makes transformation relatively easy to prepare and modify if necessary. An SQL transformation query can also be used (with some changes reflecting differences in SQL dialects) in different databases and can be executed periodically as a ‘stored procedure’.

The following conditions were used in the data transformation query:

- Select only the data for the line “automatic mode” (manual mode where the operators can interfere with a process is available for testing purposes)
- Determine the beginning and end time stamp of the robot work cycle that were to be used to compute the production cycle length
- Measure any changes for all of the energy counters (for both the electrical energy and air consumption that needed to be computed based on the cycle start and end timestamps)
- Transform the robot status bits into single enumeration

To ensure that the device state signals and energy measures were interpreted correctly, the authors also needed to develop verification methods. Besides the manual observation of the line, the behavior signals were also checked using SQL queries and RapidMiner analysis. One of the main challenges was the discovery that various signals were either probed by the PLC or reported by devices with different frequencies. For example, the power consumption peaks that are characteristic for the start of a robot’s movement could be separated from the PLC signal “robot activity” by almost one hundred milliseconds.

Figure 4 shows that there was more than a second’s difference between the change of the robot cycle counter and the change of the robot activity bit. Further examination showed that, in some cases, the robot activity bit did not change from 1 to 0 between

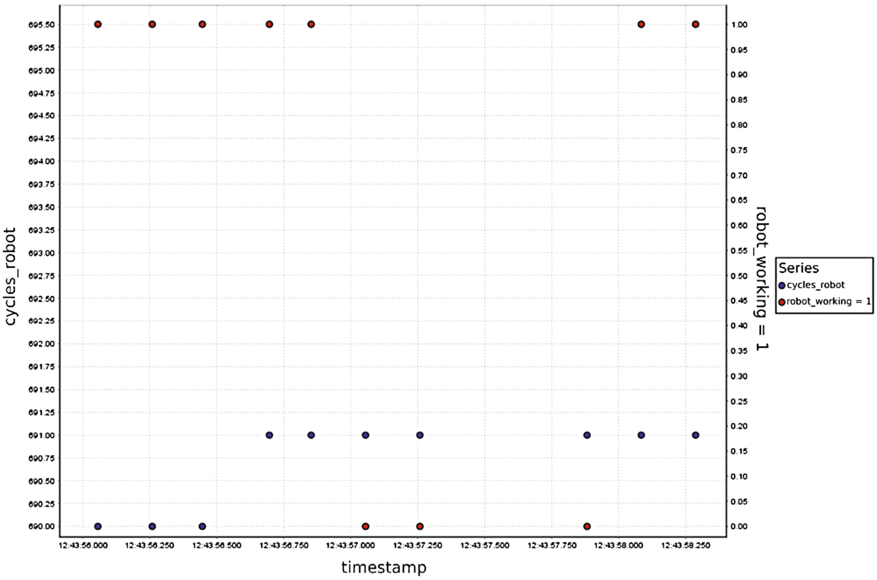

Fig. 4. Plot comparing the robot cycle counter and robot activity status.

cycles (or changed for a period shorter than the PLC scan resolution) and the only way to determine the cycle change was to look at both the cycle counter and activity bit.

Although the SQL transformation query could be applied each time an analysis was performed, in order to speed up the process, the results were stored into a comma-separated value file (CSV). While this format was used for the convenience of the analysis, for larger installations it should be replaced by a regular SQL database (energy efficiency database). Using two databases (source and transformed data) with the SQL query for the data transformation executed periodically as a database “stored procedure” would also allow old data to be removed from the source database, thus preventing it from excessive growth. Storing the transformed data greatly accelerated the execution of subsequent processes with different parameters as the transformation query took several minutes to complete and reduced the working data set for 4,000 production cycles from the gigabytes of the source snapshot data to less than 500 kB.

The initial blocks of the Rapid Miner process apply some additional filtering conditions. In general, “constant” conditions (such as the detection of station automatic mode) were used in the SQL and “dynamic” conditions (such as narrowing the time range or eliminating potentially erroneous cycles) were used in Rapid Miner processing – this permitted fast iterations with the changing processes.

The energy consumption data, which was processed without additional filtering conditions (“dynamic” conditions), exhibited several types of erroneous cycles:

- Excessive use of air
- Extremely long robot work times
- Excessive electrical energy use correlated with long robot work cycles
- Cycles with no energy consumption or no compressed air consumption

After successfully transforming, importing and filtering data, the authors looked for correlations between the production cycles and energy use. The data mining environment permits various aspects of energy efficiency to be quickly analyzed in relation to the production data.

As an example, Fig. 5 shows a plot of compressed air (blue) and electrical energy (red) use as a function of robot cycle time with a linear correlation between electrical energy use and the robot cycle length and no obvious correlation between cycle length and compressed air use.

In the case of the “OKO” line, the robot movement types could be interpreted as production variants, so it would be particularly interesting to find out whether a production variant can be determined only by its energy footprint.

An “OKO” line robot performs four main actions – it moves a new device from the input palette to one of the testing stations, it moves a new device from an additional supply drawer to one of the testing stations, it moves a device from a testing station to a laser stand (test succeeded) and it moves a device from the testing station to a drop out drawer.

Depending on the testing station being used, there are six different variants of a robot’s movements, which result in 24 program variants altogether. Theoretically, all of these variants could influence media consumption for both electrical energy and compressed air. Other interesting factors are the cycle length, type of robotic arm movement and laser use.

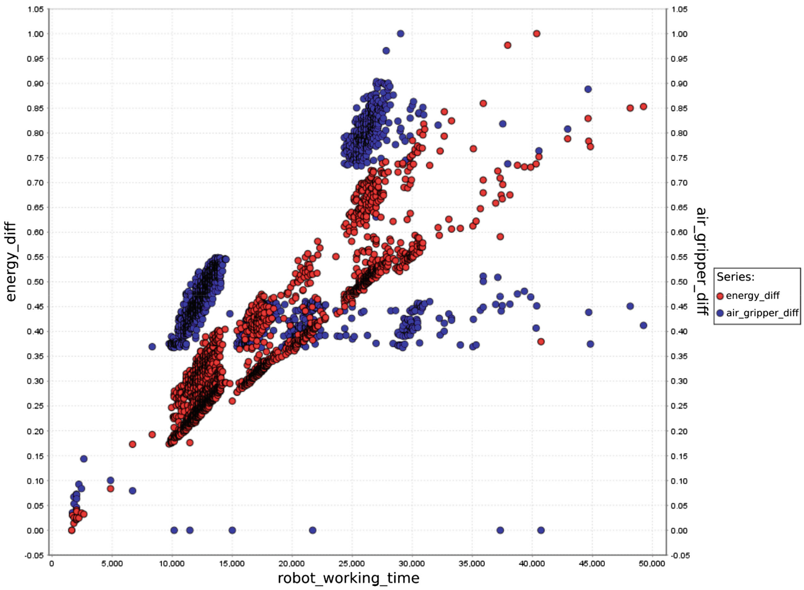

Fig. 5. Air and electrical energy consumption in function of robot working time (Color figure online)

However, the research showed that there is very little difference in energy consumption for the same robot movement type using the different testing stations. Taking this into account and considering the fact that the analysis for 24 variants significantly lowers sample size per variant, the authors decided to concentrate on four robot movements. Although the K-means algorithm was used and tested with various parameters (different distances, use of “good start values” and changes in the number of iterations), the only significant change was that the quality of clustering improved with the number of iterations that were applied. The next step was to compare clustering by energy consumption (Fig. 6) with clustering by the actual robot movement type or the production variant (Fig. 7) to determine whether this method could be successfully used to identify production variants.

Clustering is represented by the color parameter. Both compressed air and electrical energy were taken into consideration.

In both cases one cluster was identified correctly, as showing a significantly higher energy consumption (air and electrical), and this cycle was also longer. This reflects the robot movement type “test succeeded” which involves the laser engraving of elements. For the other types of robot actions, the energy footprint was not directly related to robot movement type.

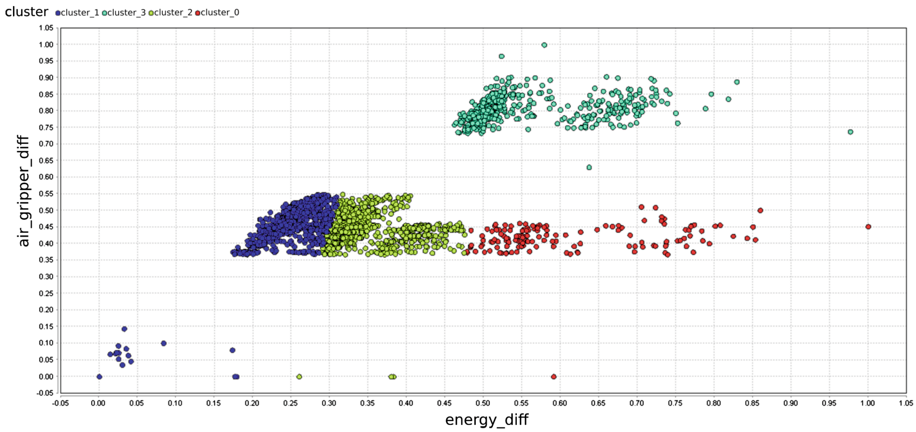

Fig. 6. K-means clustering by energy consumption (Color figure online)

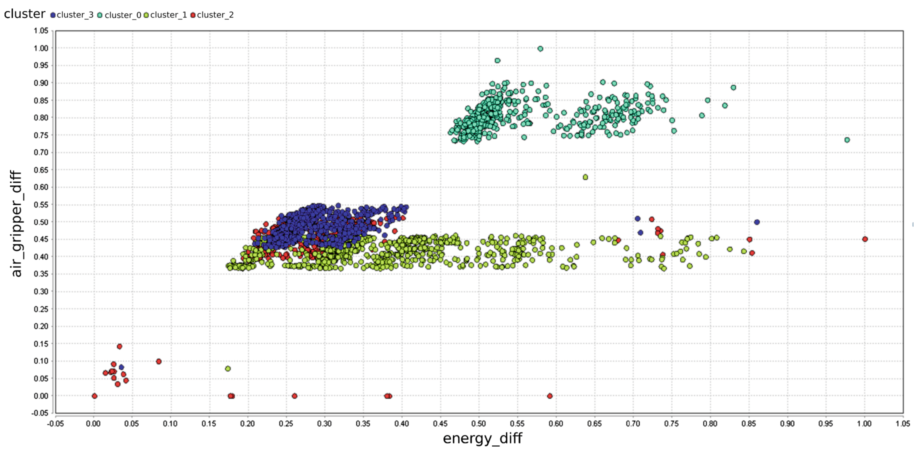

Fig. 7. K-means clustering by robot movement type (Color figure online)

5 Conclusions

In this paper, the authors have described and proposed algorithms that are dedicated to the automatic classification of the energy consumption data that is relevant to machine production cycles. The presented method is based on observing the behavior of a machine (control signals) and can be used in the case of mass-customized production. A predictive maintenance diagnostic system uses the results of the proposed clustering algorithm to automatically assess energy efficiency on the production station level energy consumption model that can be used to both present and process energy efficiency data using data mining tools. The authors applied the proposed algorithm to aggregate a large input data set that was obtained from a control system into information that was relevant to data mining algorithms. Then, data mining analyses were applied. The results were evaluated on the production line called “OKO”. It was proven that the proposed approach permits abnormalities related to energy consumption to be detected and classified in the context of the production cycle.

Acknowledgment. This work was supported by the European Union from the FP7-PEOPLE-2013-IAPP AutoUniMo project “Automotive Production Engineering Unified Perspective based on Data Mining Methods and Virtual Factory Model” (grant agreement no. 612207) and research work financed from the funds for science in the years 2016–2017, which are allocated to an international co-financed project (grant agreement no. 3491/7.PR/15/2016/2).

References

1. Peng, Y., Dong, M., Zuo, M.J.: Current status of machine prognostics in condition-based maintenance: a review. Int. J. Adv. Manuf. Technol. **50**(1), 297–313 (2010)
2. Bunse, K., Vodicka, M., Schönsleben, P., Brülhart, M., Ernst, F.O.: Integrating energy efficiency performance in production management–gap analysis between industrial needs and scientific literature. J. Clean. Prod. **19**(6), 667–679 (2011)
3. Cupek, R., Drewniak, M., Zonenberg, D.: Online energy efficiency assessment in serial production-statistical and data mining approaches. In: 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE), pp. 189–194 (2014)
4. Ziębiński, A., Świerc, S.: The VHDL implementation of reconfigurable MIPS processor. In: Cyran, K.A., Kozielski, S., Peters, J.F., Stańczyk, U., Wakulicz-Deja, A. (eds.) Man-Machine Interactions. Advances in Intelligent and Soft Computing, vol. 59, pp. 663–669. Springer, Heidelberg (2009). doi:10.1007/978-3-642-00563-3_69
5. Flak, J., Gaj, P., Tokarz, K., Wideł, S., Ziębiński, A.: Remote monitoring of geological activity of inclined regions – the concept. In: Kwiecień, A., Gaj, P., Stera, P. (eds.) CN 2009. CCIS, vol. 39, pp. 292–301. Springer, Heidelberg (2009). doi:10.1007/978-3-642-02671-3_34
6. Figueiredo, V., Rodrigues, F., Vale, Z., Gouveia, J.B.: An electric energy consumer characterization framework based on data mining techniques. IEEE Trans. Power Syst. **20**(2), 596–602 (2005)
7. Han, J., Kamber, M., Pei, J.: Data Mining: Concepts and Techniques, 3rd edn. Morgan Kaufmann, Burlington (2011)
8. Duflou, J.R., Sutherland, J.W., Dornfeld, D., Herrmann, C., Jeswiet, J., Kara, S., et al.: Towards energy and resource efficient manufacturing: a processes and systems approach. CIRP Ann. Manuf. Technol. **61**(2), 587–609 (2012)
9. Cannata, A., Karnouskos, S., Taisch, M.: Energy efficiency driven process analysis and optimization in discrete manufacturing. In: 35th Annual Conference of IEEE Industrial Electronics, IECON 2009, pp. 4449–4454. IEEE (2009)