



**HAL**  
open science

# Analysis of Manufacturing Process Sequences, Using Machine Learning on Intermediate Product States (as Process Proxy Data)

Thorsten Wuest, Christopher Irgens, Klaus-Dieter Thoben

► **To cite this version:**

Thorsten Wuest, Christopher Irgens, Klaus-Dieter Thoben. Analysis of Manufacturing Process Sequences, Using Machine Learning on Intermediate Product States (as Process Proxy Data). 19th Advances in Production Management Systems (APMS), Sep 2012, Rhodes, Greece. pp.1-8, 10.1007/978-3-642-40361-3\_1 . hal-01470614

**HAL Id: hal-01470614**

**<https://inria.hal.science/hal-01470614>**

Submitted on 17 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Analysis of Manufacturing Process Sequences, using Machine Learning on Intermediate Product States (as Process Proxy Data)

Thorsten Wuest<sup>1</sup>, Christopher Irgens<sup>2</sup>, Klaus-Dieter Thoben<sup>1</sup>

<sup>1</sup>Bremer Institut für Produktion und Logistik GmbH (BIBA), Hochschulring 20, 28359 Bremen, Germany

{wue, tho}@biba.uni-bremen.de

<sup>2</sup>University of Strathclyde, Glasgow, G1 1XQ, UK

chris.irgens@dmem.strath.ac.uk

**Abstract.** Quality and efficiency increased in importance over the last years within the manufacturing industry. To stay competitive companies are forced to constantly improve their products and processes. Today's information technology and data analysis tools are promising to further enhance the performance of modern manufacturing. In this paper, at first, the concept of the product state based view in a distributive manufacturing chain is presented, followed by a brief introduction of relations between product states along the chain. After showing that a in detail description based on cause-effect models is not economical viable today, the possibilities of using machine learning on intermediate product states to analyze the process sequence is introduced and discussed. Providing a chance to analyze large amounts of data with high dimensionality and complexity, machine learning tools combined with cluster analysis are perfectly suited for the task at hand within the product state based concept.

**Keywords:** Manufacturing processes, machine learning, product state, proxy data, quality

## 1 Introduction

Today's global manufacturing environment is increasingly becoming more complex and competitive [1]. Quality is considered a consequential key factor of success for manufacturing companies due to this development [2]. Successful manufacturing companies have to improve constantly to meet increasingly rigorous customer requirements on product and service quality [3, 4]. A number of companies have chosen to focus on their core competencies [5] and collaborate in production networks in order to meet these challenging customer requirements [6].

These developments, especially networked production, lead to an increase of complexity of a company's business processes. The quality of business processes on the other hand, has a major influence on the company's business success [7]. Thinking of manufacturing companies, the manufacturing processes are key as they directly add

value to the product and define the company's reception on the market through the product's quality [8, 9]. In recent years, the importance of looking at the complete manufacturing chain instead of isolated single process steps in order to reduce scrap and rework and thus, increase quality was slowly recognized by industry and academia [10].

All these mechanisms and developments add to the complexity modern companies face today. One lever to help these companies cope with the complexity is to take a closer look at the value adding manufacturing processes and the information/data involved. Understanding the mechanisms of the manufacturing processes and how the product changes along the whole manufacturing process and the single process steps can help to identify improvement potential and finally to realize better product and process quality. To do so, there is a need for a new perspective on the processes, focusing on the products state, its state and state characteristic change, and the process parameters involved along the manufacturing chain [11]. During the development of this product state based view, the importance of cross process relations and (inter-) dependencies surfaced. In a complex manufacturing process chain, different causes trigger various effects and the connection between cause and effect is not necessarily clear. For example, the cause can be triggered at an early stage of the process during an operation carried out by company A. The connected effect triggered at these early stages might surface at the final process step carried out by company D. The connection, especially in a networked production, is neither obvious nor easy to analyze. But in order to optimize the manufacturing process as a whole and thus the final product quality, the connection is nevertheless important.

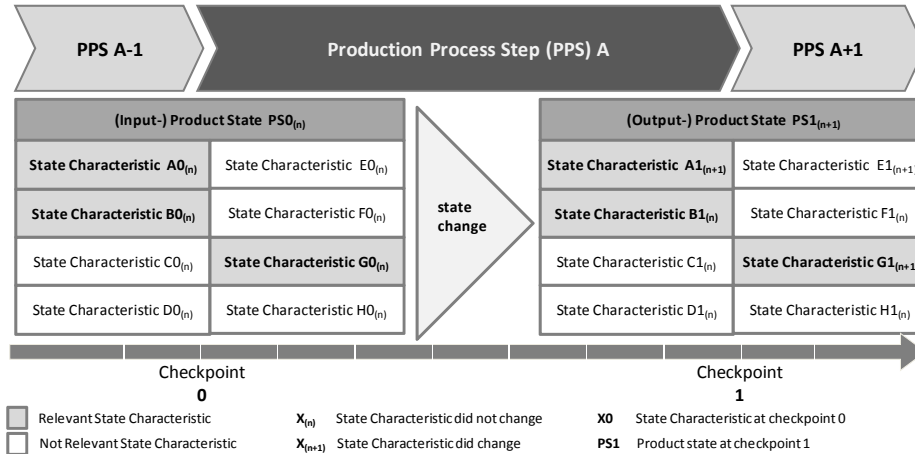
In the following chapter, the product state based view will be introduced briefly as a basic framework for the following considerations. Based on that, possible relations of states over the manufacturing process are presented and the challenges/complexity of describing such relations e.g. through cause-effect diagrams are elaborated. Following, machine learning in combination with cluster analysis is introduced as a possible solution to grasp the complexity and high dimensionality and generate results usable from a product state perspective.

## **2 Product State**

The advantage of the product state based view is the focus on the product itself to structure the information and data involved throughout the process. Using the intermediate product states as proxy data for this purpose, a product can at any time throughout its life cycle be described by its product state. Within this paper, the focus will be on the Beginning-of-Life (BOL) phase of the lifecycle, the manufacturing processes.

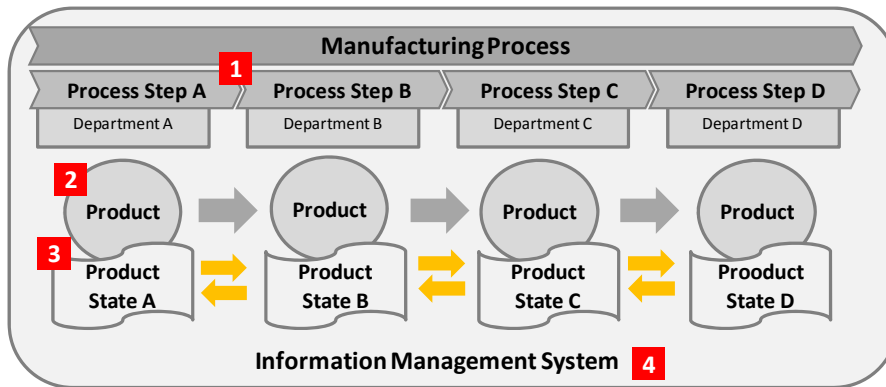
The product state describes a product at a certain time during the production process or after through a combination of (selected) state characteristics (see Fig. 1). State characteristics are definable and ascertainable measures, which can be described in a quantitative or qualitative way, e.g. weight or chemical composition of the material.

The product state changes due to external influence, for example machining or corrosion from  $t=0$  to  $t=1$  when at least one descriptive state characteristic changes [11].



**Fig. 1.** Product State Change during Manufacturing Process Step [12]

The product state can be derived theoretically at all times during the manufacturing process. In order to develop a manageable and sufficiently detailed model, the points when the product state should be derived have to be defined. The number of defined product state checkpoints during the manufacturing process affect the complexity of the model and the amount of data/information generated. It makes sense to choose at least as many check points (plus one) as there are process steps to grasp the major (intentional) product state changes.



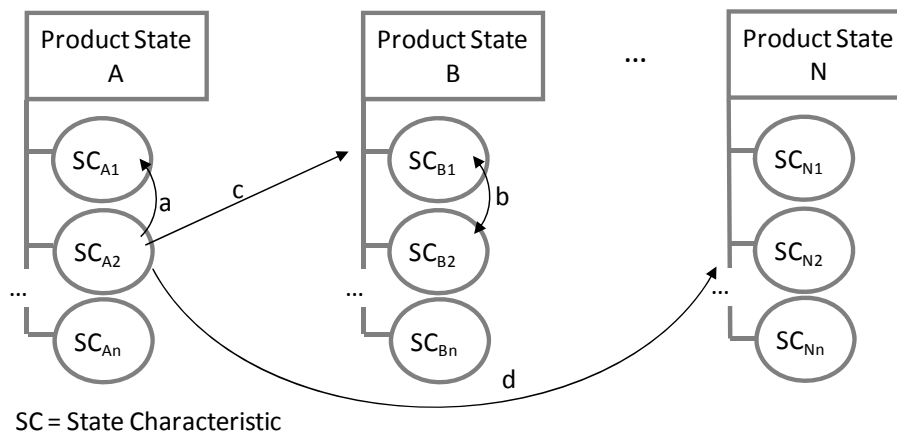
**Fig. 2.** Main Aspects of the Product State Based View [13]

Fig. 2. provides an overview on how the information and data can be structured through the product state based view. Number one in the figure indicates that the whole manufacturing process has to be considered even so not all process steps take place in one location. In that case, it is necessary to think about physical and virtual

interfaces between department or company borders. The physical product (number two) is in the center of the concept. The flow of the physical product is one directional and at each process step value is added to the product and such the product state (number three) changes by definition. In this way, the product's overall progress may be viewed as a set of concatenated 'state vectors', each vector representing one of the product states [14]. The product states have descriptive characteristics and are partly interdependent of each other as are product states over the whole manufacturing process. The surrounding information management system (number four), integrates the process (number one) and product (number two/three) information processes if needed and distributes the information to the right addressee. The distribution of the information can for example be organized through an on demand system, if the information need of the individual stakeholders in specific situations is known.

### 3 Relations of Product States

Based on the before introduced product state based view, this section will describe the relations along the manufacturing process between the different product states and state characteristics, which are helpful to provide some understanding the complexity of a modern manufacturing process chain. The term relations will be used throughout this paper compromising other terms like dependency or interdependencies to help the understanding even so the alternative terms might theoretically be more appropriate at different times. Within a complex manufacturing chain there can be a lot of different relations (see Fig. 3). They can be within a certain product state (a/b), between two directly following product states (c) or even between random product states (d). In the cases of c/d, the direction of the relation is one-directional whereas in within a product state (a/b) the direction can be reciprocal.



**Fig. 3.** Schematic Illustration of Possible Relations of State Characteristics

In theory, it would be ideal, if all relations along the manufacturing chain could be described in great detail so a profound forecast of the products final state could be

made with exact results. That way, in case of a deviation to the originally planned product state, certain adjustments could be executed accordingly to meet the customer requirements. But practically, due to the high complexity and dimensionality of the task, it is unlikely that all or even most relations can be identified and described in the needed detail. This is due to different factors e.g. too expensive, time-consuming, lack of knowledge or extensive to analyze. Therefore the chance to develop a functional and practical tool (e.g. cause-effect-diagram), based on the existing (known) product state relations, is not very likely today. A main question in this context is how relevant drivers (causes) of certain product state changes can be identified in a practical and efficient way without having to analyze the relations in such great detail. Prescriptive data processing/algorithm will find this sort of complexity hard/impossible to handle because cause-effect mechanisms may be hidden and unknown by all the relations, some of which are in themselves unknown. By basing the manufacturing process chain analysis on a 'growing' state-vector and pushing this into a machine learning process the idea is to identify the state characteristics that are likely to drive the product to success/failure.

It is thus assumed that the finished product quality, 'fitness for purpose', can be envisaged as driven by all the 'drivers' which come into play during the complete manufacturing of the product. It is clearly impractical to capture all influencing process data that is likely to contribute to the final product quality. It is thus assumed that proxy data would be useful in this context. Proxy data may be seen as two sets of parameters likely to be captured to a greater or lesser extent during the manufacturing cycle.  $S$  = the set of product states, such that

$$S = \{s_i\}, i = 0,1,\dots,m$$

and  $P$  = the set of recorded process parameters, such that

$$P = \{p_i\}, i = 1,2,\dots,m$$

where  $m$  = number of manufacturing steps and step = 0 represents the raw material/blank product input.

The proxy data sets  $S$  and  $P$  are represented by sets of vectors, matrices, such that at any stage  $k$ ,  $C_k = S_k + P_k$  represents the complete product proxy information at  $k$ .

It should be noted that  $difference(C_k)$  represents the incremental proxy state progression from stage  $k-1$  to stage  $k$ . Both  $C_k$  and  $difference(C_k)$  may be used as inputs to combined supervised and un-supervised machine learning so that analysis may be performed to determine which are the product quality driving parameters at any stage.  $C_k$  and  $difference(C_k)$  will imply a degree of cluster behaviour and the dynamics of such clustering may be used to observe the quality dynamics of the product through the manufacturing cycles. In this manner it may be possible to capture cause-effect mechanisms that influence the product quality through observing the product state progression in conjunction with the practically recorded process information.

## 4 Machine Learning

The initial discussion points above are seen as pertinent to a number of recent methodologies used in handling large multi-variate data sets. Traditional methods and techniques may be unsuitable due to high dimensionality and heterogeneous, distributed nature of the data. It is therefore considered relevant to draw on ideas from machine learning/AI, pattern recognition, statistics, and database systems.

**Cluster analysis** (CA) classifies unknown groups while discriminator function analysis classifies known groups. CA allows many choices about the nature of the algorithm for combining groups. Each choice may result in a different grouping structure.

In short, clustering analysis is about defining groups of data matching certain requirements as e.g. given a set of data points, each having a set of attributes, and a similarity measure among them, find clusters (see Fig. 4).

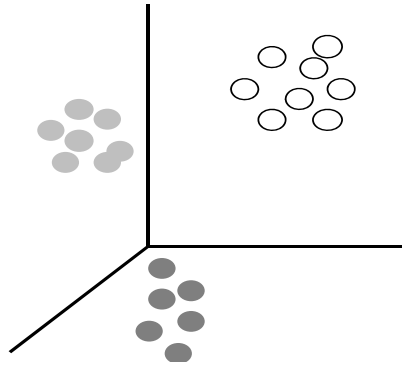


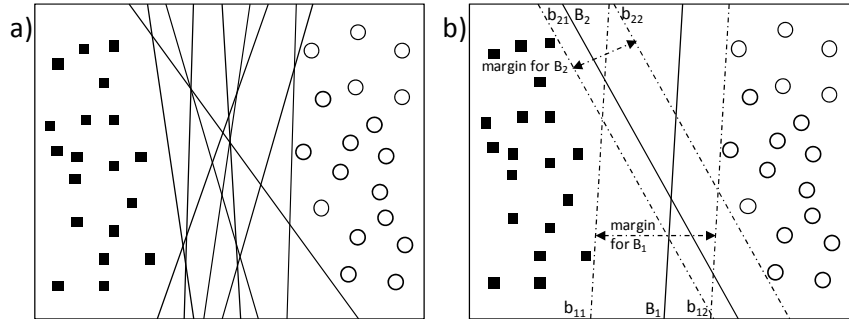
Fig. 4. Examples of Data to be Clustered [16]

When trying to establish the characteristics of a process state or regime, a clear division between the different regimes/states is required. Therefore some form of classification method would be useful.

**Support Vector Machine** (SVM) has received considerable attention. This classification technique has its roots in statistical learning theory and has shown promising empirical results in a number of practical applications and works very well with high-dimensional data. Another unique aspect of this approach is that it represents the decision boundary using a subset of the training examples, known as the **support vectors**.

The basic idea behind SVM, is the concept of a maximal margin hyper plane. A linear SVM can be trained explicitly to look for this type of hyper plane in linearly separable data, the method can be extended to non-linearly separable data.

Fig. 5 shows a data set containing examples that belong to two different classes, represented as squares and circles. The data set is also linearly separable; i.e., there is a hyper plane such that all the squares reside on one side of the hyper plane and all the circles reside on the other side. Inevitably such clear cut results are not always available in real applications and suitable compromise solutions are used in order to allow a certain amount of mis-classification.



**Fig. 5.** Possible Decision Boundaries for a Linear Separable Data Set b) Margin for Decision Boundary [16]

There are infinitely many hyper-planes possible. Although their training errors may be zero, there is no guarantee that the hyper-planes will perform equally well on previously unseen examples. The classifier must choose one of these hyper-planes to represent its decision boundary, based on how well they are expected to perform on test examples. The problem with SVM classifiers is, being able to establish relevant training/learning data for the case at hand so that the computation of the support vectors can be completed. Given that most process control and analysis task display a degree of dynamism, the use of SVM is not immediately clear for such applications as it is likely that the generation of support vectors will have to be done frequently and without knowledge of detailed performance results of the process.

Through an application of suitable tools like Cluster-Analysis combined with Support Vector Machines, it is considered likely to determine which process parameters are the product state drivers based on process or even proxy data. This way, the relevant information for adjusting the manufacturing process to reach the customers quality requirements of the individual product can be derived and put to use.

## 5 Conclusion

The paper presented an approach to derive information of the manufacturing process sequence using intermediate product states using a combination of cluster analysis and machine learning. This approach is capable of handling the complexity and high dimensionality of modern production whereas describing the actual relations between state characteristics along the manufacturing chain (through e.g. cause-effect models) soon reaches its limits of being applicable efficiently.

The introduced approach seems well suited for the tasks within the product state based view. The perspective on the product by its time-dependent state over the manufacturing process corresponds directly with the requirements of machine learning tools. Thus, the time and effort to match the (product state) data to the requirements of the machine learning tools will be within an acceptable range and can possibly even be automated in a later step.



In a next step the theoretically presented approach will be put to the test with manufacturing process data and an in depth analyses and discussion of the results will offer the chance to evaluate its prospects in an industrial application to increase quality and reduce rework and scrap.

**Acknowledgement.** The authors would like to thank the “Deutsche Forschungsgemeinschaft” for financial support via the funded project “Informationssystem für werkstoffwissenschaftliche Forschungsdaten” (InfoSys).

## References

1. Levitt, T., “The globalization of markets”, Harvard Business Review, 1983(May/June), pp. 92-102, 1983.
2. Enderwick, P., Globalization and Labor. New York: Chelsea House Publications, 2005.
3. Kovacic, M. & Sarler, B., “Application of the Genetic Programming for Increasing the Soft Annealing Productivity in Steel industry”, Materials and Manufacturing Processes, vol. 24, pp. 369-374, 2009.
4. Ellram, L.M. & Krause D.R., “Supplier Partnerships in Manufacturing Versus Non-Manufacturing Firms”, International Journal of Logistics Management, vol. 5, no. 1, pp.43-54, 1994.
5. Hamel, G. & Prahalad, C.K., “The Core Competence of the Corporation”, Harvard Business Review, 1990(33), pp. 79-91, 1990.
6. Seifert, M., Unterstützung der Konsortialbildung in Virtuellen Organisationen durch perspektives Performance Measurement. Dissertation Universität Bremen, 2007.
7. Linß, G., Qualitätsmanagement für Ingenieure. München/Wien: Hanser Verlag, 2002.
8. Brinksmeier, E., Prozeß- und Werkstückqualität in der Feinbearbeitung. Fortschritt-Berichte VDI Reihe2: Fertigungstechnik Nr. 234.Düsseldorf: VDI Verlag, 1991.
9. Jacob, J. & Petrick, K., Qualitätsmanagement und Normung, pp. 101-121. In: Schmitt, R. & Pfeifer, T. [eds.]. Masing Handbuch Qualitätsmanagement. München: Carl Hanser Verlag, 2007.
10. Zoch, H.-W. & Lübben, Th., Verzugsbeherrschung - Systemorientierter Ansatz als wesentliche Voraussetzung für den Erfolg. Stahl Strukturen - Industrie-, Forschungs-, Mikro- und Bauteilstrukturen - 26. Aachener Stahlkolloquium, Verlagshaus Mainz, 2011
11. Wuest, T., Klein, D., Thoben, K.-D., “State of steel products in industrial production processes”. Procedia Engineering, 10(2011), pp. 2220-2225, 2011.
12. Wuest, T., Klein, D., Seifert, M., Thoben, K.-D., “Method to Describe Interdependencies of State Characteristics Related to Distortion. *Mat.-wiss. u. Werkstofftech.*, 1-2(43), p. 186-191, 2012.
13. Wuest, T. & Thoben, K.-D. “Exploitation of Material Property Potentials to Reduce Rare Raw Material Waste - A Product State Based Concept for Manufacturing Process Improvement.” *Journal of Mining World Express (MWE)*, 1(1), 2012.
14. Irgens, C., Wuest, T, Thoben, K.-D., Product state based view and machine learning: A suitable approach to increase quality? Proceedings of the 14<sup>th</sup> IFAC Symposium on Information Control Problems in Manufacturing, May 23-25, 2012, Bucharest, Romania, 2012.
15. Mirkin, B.. Clustering for Data Mining, A Data Recovery Approach. London NY: Chapman & Hall/CRC, 2005.
16. Pang-Ning T., Steinbach, M. & Kumar, V.. Introduction to Data Mining. Old Tappan: Pearson/Addison Wesley, 2005.