An Overview of Useful Data and Analyzing Techniques for Improved Multivariate Diagnostics and Prognostics in Condition-Based Maintenance

Carolin Wagner[1], Philipp Saalmann[1], and Bernd Hellingrath[1]

[1]*Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany*
{wagner; saalmann; hellingrath} @ercis.uni-muenster.de

ABSTRACT

The reliability of production machines gains in importance in today's optimized and highly productive business environments. Unexpected machine breakdowns do not only lead to loss of production time and production outages but also to diminishing customer satisfaction due to deterioration in quality and declining availability of products. The condition-based maintenance (CBM) strategy aims at preventing these machine breakdowns through real-time monitoring of machine conditions. Sensor data are collected and analyzed using diagnostic and prognostic approaches to identify the type of fault and the remaining useful life. Identifying the reasons and time of breakdowns fosters improved planning of maintenance and spare parts demand, leading to higher machine reliability. In general, machine sensor data are regarded as a useful source of information to assess the machine's operating condition. However, in some specific cases, the machine sensors lack the ability to correctly represent the health of the machine or the specific component under consideration. Therefore, additional information by further available data is required to improve diagnostic and prognostic techniques for more accurate and precise analysis. Current research focuses on the analysis of sensor data for condition-based maintenance, while other data like the operating history and environment temperature have only been considered to a limited extend so far. Hence, this paper gives an overview on potential data for machine health assessment and remaining useful life prediction in condition-based maintenance. Furthermore, corresponding approaches and techniques for fault diagnostics and prognostics are presented targeting the analysis of individual data sources as well as of multivariate settings featuring multiple integrated data sources.

1. INTRODUCTION

Unexpected machine breakdowns are one of the primary concerns in production companies today. Due to the drawbacks of reactive and preventive maintenance strategies with regard to machine availability and costs, current research focuses on the application of condition-based maintenance. Condition-based maintenance (CBM) is characterized by the monitoring and analysis of machine conditions through Intelligent Maintenance Systems (IMS), which facilitate the surveillance of the machine status based on the machine's sensors. These sensor data provide information to identify and forecast device failures, resulting in more precise maintenance and spare parts requirements.

The success of the condition-based maintenance strategy is mostly reliant on the available data and the techniques used for evaluation. Fault diagnostics and prognostics are almost exclusively based on sensors that are installed in the specific machine. Information gained from these sensors though is limited to the observed objects, i.e. selected machine components, and thus cannot capture the complete condition of the machine. However, companies have access to huge amounts of additionally influential data such as operating and maintenance history, production environment conditions and expert knowledge of technicians, which should be included in the failure detection and remaining useful life techniques for improving prediction accuracy.

Most of the research in the field of condition-based maintenance has focused on the development of techniques that only use machine sensor data for assessing the machine's condition – disregarding additional data sources. Improvement of existing methods by combining multiple techniques has been addressed by many researchers (Peng, Dong & Zuo, 2010). Moreover, some researchers have already suggested further data sources to be considered for condition monitoring, e.g. the operating condition (Li, Zhan & Li, 2014). However, a comprehensive overview of possible data sources and techniques has not been published yet.

This paper presents a summary of useful data sources and corresponding techniques for diagnostics and prognostics in condition-based maintenance. Based on the identified data sources, suitable approaches and techniques for machine health assessment and remaining useful life prediction are presented. They cover the analysis of individual data sources as well as of multivariate settings featuring multiple integrated data sources. The structure of the paper is as follows: chapter 2 provides an overview of condition-based maintenance, including the two main aspects fault diagnostics and remaining useful life prognostics. In chapter 3 a literature review on exiting data classifications is provided. Based on these results different groups of data sources are introduced in detail in chapter 4. This is followed by an investigation of corresponding groups of diagnostic and prognostic techniques in chapter 5.

2. CONDITION-BASED MAINTENANCE

The maintenance strategy *condition-based maintenance* is based on the supervision of machines using sensory information. It enables the assessment of the current machine condition in real-time as well as the prognosis of future machine degradation in order to increase reliability of machines and decrease the probability of machine breakdown. In condition-based maintenance, the machine is either monitored continuously or at regular intervals.

There are two main elements, which are important for condition-based maintenance:

- *Fault diagnostics* refers to the analysis of condition data in real-time in order to detect failures as soon as they become visible in the condition data. Techniques used for fault diagnostics aim at identifying the location of failure occurrence and the type of fault (Dai & Gao, 2013) in order to schedule the required maintenance activity and provide the right spare part if necessary.
- The goal of *prognostics* is the estimation of the remaining useful life of a machine. The remaining useful life (RUL) is defined as the duration between the current time and the time of machine breakdown. The terms prognostics, remaining useful life and time-to-failure are often used simultaneously. Prognostics include current machine health estimation, health propagation and failure threshold determination (Saxena, Sankararaman & Goebel, 2014).

Methods for condition-based maintenance can be grouped into three classes: model-based, knowledge-based and data-driven models (Dai & Gao, 2014; Peng et al., 2010). For model-based approaches, domain experts develop a mathematical model considering the physical processes. Knowledge-based model building is based on historical data in combination with domain experts. Lastly, in data-driven approaches the model is built from the data only using statistical and artificial intelligence learning techniques (Peng et al., 2010).

In this paper, the general term *machine* is used to refer to a distinct part of a system for which condition-based maintenance is implemented. Since machines can get very complex, CBM is often used only for smaller units or components of the considered machine. Identified data sources and techniques can however be directly transferred to machine sub-units or components.

3. EXISTING DATA CLASSIFICATION

Research in the field of condition-based maintenance has primarily focused on the development of methods for fault diagnostics and prognostics. No paper has been identified which provides a detailed presentation of useful data sources in condition-based maintenance. Nevertheless, some publications already defined different groups of data in order to distinguish between the groups as part of their work.

A general data classification is presented by Jardine, Lin and Banjevic (2006). Data are classified into condition monitoring data and event data. The former includes the measurable data related to the health condition of the machine, while the latter covers all information about past failures and the maintenance history of the machine. In order to review existing methods for data analysis in condition-based maintenance, a further distinction between three different types of condition monitoring data is provided. Those types include the value type (single value), the waveform type (time series) and the multidimension type (multidimensional data). Mobley and Keith (2002) identify and analyze in detail the different possible machine condition sensory data, in their work referred to as predictive maintenance techniques. On the other hand, Ghodrati (2005) defined several influential factors specifying the operating condition in order to improve the performance of the proportional hazard model. They have been able to show that the consideration of working environment and operator skill have a significant influence on the model accuracy. A similar approach is suggested by Tsang, Yeung, Jardine, and Leung (2006). They present a data structure including failure data, condition data, maintenance action data and installation data, which should be used for proportional hazard. A detailed overview of their data structure is shown in Figure 1. A data acquisition procedure to accurately deal with data quality problems is defined by Moss (1991). In this procedure, different data types are listed to provide guidelines for data to be used during the reliability analysis of machinery. The three data types are inventory data (e.g. item identification, environmental parameters), failure data (e.g. time interval, operational hours) and maintainability data (e.g. man hours, active repair times).

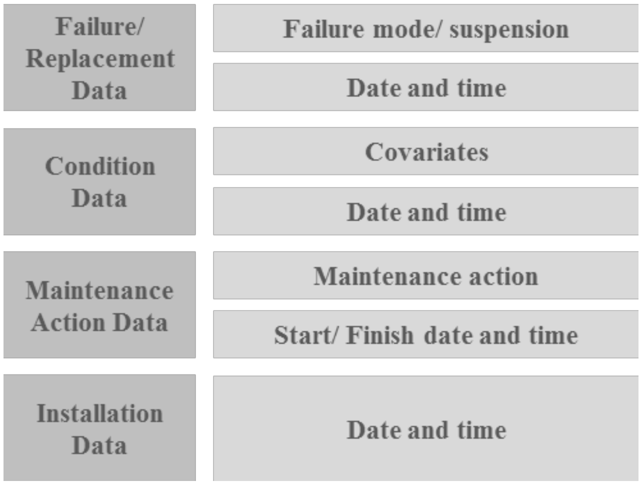

Figure 1. Data classification by Tsang et al. (2006)

These publications present several attempts to perform data classification which can be used for further analysis. However, none of the presented publications showed a thoroughly investigation of possible data sources for condition-based maintenance. The following data classification is based on these approaches and additionally extended in order to build a more comprehensive review of useful data sources for fault diagnostics and prognostics.

4. DATA GROUPS AND DATA SOURCES

Data gained from sensors are the basis for data-driven, condition-based maintenance. Classification and learning algorithms use historical and real-time data in order to diagnose the current machine health status and estimate future machine degradation. The accuracy of those methods can be improved by including further data in the analysis.

One additional crucial factor is data quality, which is especially relevant in real-world settings. Since the quality of the final prediction is highly dependent on the quality of the available data, good data quality is essential for condition-based maintenance (Saxena et al., 2014). Data are either acquired automatically or entered manually (Jardine et al., 2006). The latter is particular error-prone and thus should be carefully preprocessed before usage. In this study automatic acquisition of data is focused due to the demanded data quality. In general, data are assumed to be cleaned beforehand and has a high resulting quality.

Four groups of data sources are identified by analyzing data sources from literature with regard to their origin and scope. An overview of these groups with related data sources is presented in Figure 2.

- The most important data source is *machine condition data*. This group refers to sensory information measuring specific inner-machine characteristics. Through machine monitoring, e.g. via vibration monitoring or oil analysis, machine degradation and fault progression can be observed. Machine condition data are essential for condition-based maintenance since real-time information about the current health status of the machine is provided.
- Machine condition data can be enhanced with available information about the machine in order to obtain a better and more profound picture of the current health of the machine. This is especially important since the machine degradation can vary along the lifecycle of a machine (Voisin, Medina-Oliva, Monnin, Leger & Iung, 2013). Therefore, *machine-related information* including age, the current and historic workload, and past maintenance activities can additionally be integrated for analysis in order to improve the capabilities of diagnostic and especially prognostic methods.
- Besides considering the individual machine, *external factors* can have an implication on the degradation behavior. Machines are exposed to specific production environments. Machines working under e.g. high humidity and high temperature will degrade faster and failures are more likely.
- The group *supplementary sources* refers to acquired data, which has no direct effect on the machine. For example, employees working with the machine have a good knowledge about the current machine condition, which might not be captured from sensors. Furthermore, information from similar machines (within the same fleet) can give insights into future degradation behavior.

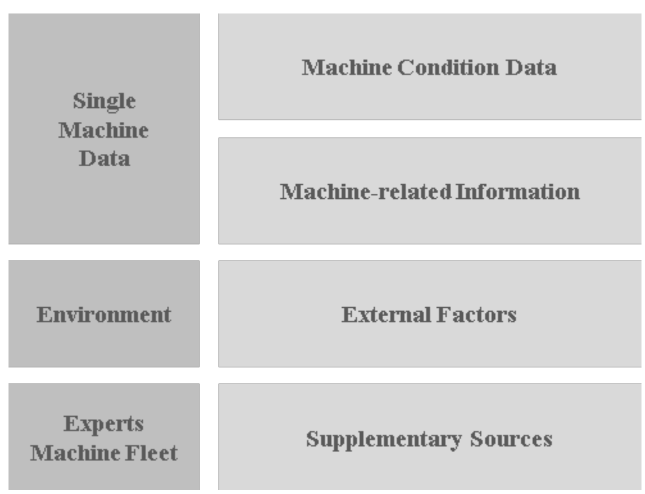

Figure 2. Overview of data sources for CBM

4.1. Machine Condition Data

Machine condition data is the core data source for condition-based maintenance. Sensors, which are attached to specific parts of the machine, are able to measure inner-machine characteristics. These data provide the information base for diagnostics and prognostics.

There are different data acquisition approaches for machine condition data collection. Well-known techniques include vibration monitoring, thermography and oil analysis (tribology) (Mobley & Keith, 2002). Out of these techniques, *vibration monitoring* is the most important data acquisition tool for monitoring the mechanical conditions of machinery. Vibration sensors are attached to machine parts or components and measure amplitudes of the oscillation (Mobley & Keith, 2002. Another data acquisition approach is *infrared thermography*, which measures the emission of infrared energy using i.e. infrared thermometers or imaging systems (Mobley & Keith, 2002). *Oil analysis*, also known under the term tribology, refers to bearing-lubrication structures of machinery. The two approaches for oil analysis are either lubricating oil analysis or wear particle analysis (Mobley & Keith, 2002). In the first case, oil quality is checked frequently and exchanged if required. The latter provides direct information about the wearing condition of the component by assessing particle shape, composition, size, and quantity (Mobley & Keith, 2002). Other approaches to collect machine condition data include electrical testing (Mobley & Keith, 2002), sound analysis (Bengtsson, Olsson, Funk & Jackson, 2004), or the application of ultrasonic sensors (Jardine et al., 2006).

In many cases, several sensors monitor the machine condition in order to include different sensor positions and data acquisition techniques. The collection of data from multiple sensors provides two main advantages over single source data analysis. Combining same-source signal data can reduce uncertainty with regard to erroneous measurements. In addition, the accuracy of the determination of the machine health condition is increased using multiple types of sensors, which can measure different aspects of the system (Hall & Llinas, 1997).

4.2. Machine-Related Information

The degradation of machines depends on different characteristics of the machine and might change in the course of the machine lifecycle (Bonissone & Varma, 2005). It is therefore important to include machine-related information into time-to-failure estimations. Machine-related information includes all information about the machine, which are not captured by installed sensors.

4.2.1. Master Data

Machine master data contain the basic information about the machine, which includes among others the age and the components list of the machine. The former is particular interesting for condition-based maintenance. In general, the reliability of machines decreases over the course of the machine's life. Degradation of the machine is not necessarily visible in the early stages. Even though the remaining useful life decreases, this change does not become apparent in the machine health. Most diagnostic and prognostic algorithms however are exclusively based on the machine's current health index without consideration of its age. This issue is not only crucial for early machine stages, but is also valid for later phases.

4.2.2. Operating History

The operating history refers to the historical and current load and speed of the machine's mode of operation. The utilization (load and speed) of machines has high impact on the machine degradation. Over time, machines are exposed to different workloads, which impact the machine condition (Liao & Lee, 2009). Machines under high workload and speed will degrade faster and are more likely to run into failures. Additionally, the total operating time of the machine is captured by the operating history, which similar to the age is an additional good indicator for remaining useful time estimations.

4.2.3. Maintenance History

Machines are usually not operated until breakdown, but maintained periodically. Depending on the maintenance strategy or legal requirements this could either be after a pre-defined time frame or when maintenance demand is indicated by the machine condition. All maintenance actions on a specific machine are stored in a so-called maintenance history. This does not apply for components, which are directly replaced and not repaired. Several types of maintenance actions exist. Maintenance include i.e. general overhaul, component replacement, and partial repair (Ghodrati, 2005). Performing maintenance on a machine increases life expectancy however, initial machine conditions will not be reached. Poor maintenance or incorrect spare parts on the other hand have negative impacts on the machine's condition (Ghodrati, 2005). The quality of machine maintenance has a great impact on condition-based maintenance. The maintenance history could therefore be used to improve the accuracy of diagnostic and prognostic techniques.

4.3. External Factors

External factors include the acquisition of data, which have a direct impact on the machine but are not originating from the machine itself. The external factors of machines are mainly determined by the *production environment* to which the machine is exposed. For machines operating in a non-optimal production environment, remaining useful life is decreased and failures are more probable to occur. Two different influential factors can be distinguished: the *climatic* *condition* and the *physical environment* (Ghodrati, 2005). The climatic condition is influenced by the weather, which impacts the microclimate in the production area. In order to quantify the climatic condition, temperature and humidity of the environment are good indicators (Ghodrati, 2005). Weather forecasts can additionally be used to predict future climatic conditions. In order to assess long-term effects on the remaining useful life, seasonal changes can be anticipated. The physical environment on the other hand is influenced i.e. by dust, smoke, fumes and corrosive agents (Ghodrati, 2005).

4.4. Supplementary Sources

Supplementary sources can provide additional information about the machine, which do not originate from the machine itself and have no effects on the machine. Two different sources have been identified in the literature: expert knowledge and fleet knowledge.

4.4.1. Expert Knowledge

Expert knowledge refers to the expertise gained by humans who are in contact with the machine, e.g. maintenance personnel or machine operator. Humans can acquire information about the machine i.e. by visual inspection, by paying attention to sounds, by touching the machine or by transferring experiences gained from other machines. The problem with human knowledge is that most knowledge is available only implicitly and lost when the expert leaves the company. In order to access this knowledge, information needs to be acquired from the expert and transferred in a machine-readable format. The acquisition and processing of expert knowledge is particularly important in cases where the machine fails and knowledge about the cause cannot be extracted from existing data.

4.4.2. Fleet Knowledge

The accuracy of diagnostic and prognostic approaches can additionally be enhanced using information and knowledge obtained from machines within the same fleet (population-based analysis). In this context, a fleet is defined as a set of machines, which are grouped with regard to some characteristics. While previously mentioned data sources focus on different aspects that might impact the observed machine, the fleet-based approach considers multiple machines to increase the amount of data available for health assessment and remaining useful life calculation. These additional data can be used to detect similarities within the degradation behavior of different machines. Prediction of future machine health status can therefore be supported by data obtained from machines working in the same operational context and at a similar stage of the lifecycle (Voisin et al., 2013). Besides reducing uncertainties through increased data availability (Pecht & Jaai, 2010; Medina-Oliva, Voisin, Monnin, Peysson, Léger, 2012) the long time prediction accuracy can be improved with respect to information on past machine degradation behavior (Liu, Djurdjanovic, Casoetto & Lee, 2007).

4.5. Benefits of Data Sources

Different data sources for improving condition-based maintenance have been presented in this chapter. A summary of these data sources is provided in Table 1. As outlined, machine condition data is the key element for diagnostics and prognostics. Nevertheless, analysis of data should not only focus on inner-machine characteristics measured by sensor data but also include additional data sources as presented above. Two major influence factors of the machine health are the operating workload and speed of the machine as well as the production environment. Those influence factors can in certain cases contribute significantly to the accuracy of prognostic and diagnostic techniques. In case, data for condition-based maintenance are acquired for several identical machines, the fleet knowledge also provides an important data source for prognostics. Expert knowledge, general master data as well as the maintenance history are data sources, which will probably add only little additional information, since most is already captured by the sensory information, however can provide good indicators if available data cannot measure the health of the machine precisely.

Table 1. Overview of available data for CBM

Data Group	Data Source
Machine Condition	Inner-Machine Characteristics
Machine-related Information	Master Data
	Operating History
	Maintenance History
External Factors	Production Environment
Supplementary Sources	Expert Knowledge
	Fleet Knowledge

5. INTEGRATION TECHNIQUES

In order to improve diagnostic and prognostic accuracy for condition-base maintenance, data from different sources has to be integrated with inner-machine condition data. In sub-chapter 5.1 well-known techniques for condition-based maintenance are presented, which are based on machine condition sensor data. In the subsequent sub-chapters, several integration techniques are pointed out which combine the respective data source with inner machine condition data. Techniques for machine-related information and external factors are identical and are therefore described combined in one sub-chapter. This study will mainly focus on nonprobabilistic methods, since they are considered suitable for multivariate data analysis.

5.1. Machine Condition Data

There are various methods available to analyze inner-machine sensor data. Cecati (2015), Sikorska, Hodkiewicz and Ma (2011) and Si, Wang, Hu and Zhou (2011) provide detailed overviews of data-driven techniques for diagnostics and prognostics using machine condition data. In general, diagnostic methods are known as supervised classification methods from the field of pattern recognition and prognostic methods are referred to forecasting techniques, especially trend extrapolation. However, the amount of possible methods is immense, and not only limited to these two areas.

Techniques for data-driven fault diagnostics can be grouped into statistical and non-statistical methods depending on how quantitative information is obtained (Cecati, 2015). Specific statistical methods include principal component analysis, partial least squares, independent component analysis and support vector machines (Cecati, 2015). On the other hand, well-established non-statistical methods are neural networks and fuzzy logic (Cecati, 2015).

Data-driven prognostic methods for condition inner-machine sensor data on the other hand can be classified into three types of methods: stochastic (probabilistic) and statistical life expectancy models as well as artificial neural networks (Sikorska et al. (2011). Examples for specific techniques are Markov models (stochastic), auto-regressive integrated moving average model (statistical) and recurrent neural networks (neural networks).

In order to analyze the data streaming from multiple sensors, multi-sensor data fusion techniques are applied. Those techniques are classified into data-level fusion, feature-level fusion, or decision-level fusion (Jardine et al., 2006).

5.2. Machine-Related Information and External Factors

Machine-related information and external factors are either single-value, multi-value (e.g. master data, temperature, humidity) or event data (maintenance history). In order to combine this additional information with inner-machine condition data, techniques are applied, which can handle more than one input variable.

Two important groups of techniques for multiple input variables have been identified which have frequently been applied to condition-based maintenance, namely machine learning techniques and survival models. While the former group presents generic supervised learning techniques, which can be used for both diagnostic and prognostic problems using multiple input variables, the latter group is specifically designed for problems for which the time until occurrence of an event is estimated (Rodríguez, 2010)

Machine learning methods are usually very flexible with respect to the number of input vectors and specific characteristics of the problem. They perform well when huge amounts of data are available and are highly adaptable to the problem at hand. Due to this flexibility and adaptability, machine-learning methods have been widely applied in condition-based maintenance (both for inner-machine condition data and with multiple inputs). Common machine learning techniques include neural networks and support vector machines. However, most machine learning techniques are black-box methods. Therefore, it is often difficult to obtain more information from the model other than the current machine health status or the remaining useful machine life (Li et al., 2014).

Survival models present the second group of well-known techniques. They are characterized by waiting time estimation including different covariates in the analysis (Rodrígues, 2007). These characteristics can be directly matched to the requirements of remaining useful life estimation with multiple input variables. A common example for survival methods is proportional hazard models (Jardine et al., 2007). Influencing factors are included as covariates to the prognostic model, which are able to increase or decrease the observed degradation (hazard).

Besides these two groups, there are other categories including Bayesian networks and multivariate trend exploration, which are capable of learning multiple input variables. Yet, machine learning and survival methods have been identified as the two main groups of methods for condition-based maintenance.

5.3. Supplementary Sources

Techniques to integrate supplementary sources with inner-machine condition data are strongly dependent on the source and the available data. Therefore, each data source needs to be discussed separately.

Human domain knowledge can be captured by experts. This knowledge can then be used to build an expert system. Expert systems have a long history in diagnostic and prognostic applications (Peng et al., 2010). Domain knowledge can be gathered and transformed into valuable information using a set of rules. The resulting knowledge base is used for the inference of knowledge. Combining expert systems with other artificial intelligent methods can improve the accuracy of condition-based maintenance techniques (Angeli, 2010). In order to include expert knowledge into diagnostic and prognostic techniques, a data-driven condition-based maintenance technique could be enhanced using rule-based expert systems. Other techniques which are often used to capture and integrate expert knowledge are fuzzy sets (Peng, Dong & Zuo, 2010) and Bayesian networks (Sikorska et al, 2011).

As an additional data source, *fleet-based information* can be used to improve remaining useful life estimation (prognostics). Three different types of fleets can be distinguished with respect to their composition of machines. They either comprise identical/ homogeneous machines, similar machines or heterogeneous machines (Medina-Oliva et al., 2012). Approaches to consider fleet information are developed with regard to one specific type of fleet. While approaches to deal with homogenous and heterogeneous fleets have been developed in research, similar machines have rarely been addressed (Medina-Oliva et al., 2012).

Data obtained within *homogeneous fleets* can directly be used to decrease uncertainty through higher data availability. Three interesting approaches for homogenous fleets have been identified. Liu et al. (2007) and Wang et al. (2008) use distance measurements with regard to the machine degradation in order to identify the similarity between other machines of the fleet. Future degradation of the machine is assessed by matching the degradation behavior of machines, which show the highest similarities. In the procedure developed by Turin, Subbiah, Leone & Cristaldi (2015), the slope of the degradation during a pre-defined time frame is calculated for each machine in the fleet. Remaining useful life of the machine is then estimated via simulations that randomly sample the degradation rate of other machines in the same time interval.

Heterogeneous fleets, on the other hand, consist of heterogeneous units, which have some similarities in common (Voisin et al, 2013). Thus, data from all machines in the fleet cannot directly be transferred to the considered machine. By matching the technical, dysfunctional, operational, service, and application context, similarities between machines can for example be identified (Monnin, Abichou, Voisin, & Mozzati, 2011; Medina-Oliva et al., 2012).

5.4. Summary of Techniques

Different groups of techniques have been presented in this chapter, which can be applied for diagnostics and prognostics. In general, diagnostic approaches use classification techniques and prognostics apply forecasting techniques. An overview of the different identified groups of techniques is shown in Table 2.

Most of the techniques used for analyzing machine condition data (first data group) are only able to include one single input variable. In order to include additional data in the analysis and by thus improving the accuracy, methods, which can deal with multiple input variables, have to be considered. A well-known group of techniques, which is very flexible both to the number of input variables and the specific problem at hand, are machine-learning algorithms. However, machine-learning techniques require a lot of historical data and present black-box approaches. Survival models on the other hand require little historical data and are simple to comprehend. The drawback of survival models, however, is their limitation to prognostics only as well as the additional effort to build the model. Expert Systems and similarity measures are two further approaches to include specific supplementary data sources to improve techniques for condition-based maintenance, but are resource-intensive in their creation.

Table 2. Overview of different groups of techniques

Group of Techniques	Data Group	Purpose
Pattern Recognition	Machine Condition	Diagnostics
Forecasting	Machine Condition	Prognostics
Trend Extraploation	Machine Condition/ Machine-related Information/ External Factors	Prognostics
Machine Learning	Machine Condition/ Machine-related Information/ External Factors	Both
Survival Models	Machine Condition/ Machine-related Information/ External Factors	Prognostics
Expert System	Supplementary Sources (Expert Knowledge)	Both
Similarity Measures	Supplementary Sources (Fleet Knowledge)	Prognostics

6. CONCLUSION

This study provides an overview of data sources, which could be used to improve fault diagnostic and prognostic algorithms for condition-based maintenance. Four different groups of data sources are identified, namely machine condition data, machine-related information, external factors, and supplementary sources. For each group concrete sources are considered and evaluated in more detail. Most promising further data sources are the production environment (measured by e.g. the humidity) as well as the machine load and speed. Information obtained from similar machines can additionally benefit the accuracy of the remaining useful life calculation. In a second step, different groups of techniques for machine condition data analysis and the integration of multiple data sources for diagnostics and prognostics are presented. Benefits and disadvantages are shortly outlined. In case, huge amount of data are available machine-learning techniques are well suited. If little historical data are available, survival models depict promising methods, which enable handling multiple input variables.

Future research should target in-depth research and evaluation of the presented techniques for each group of data sources. The performance of specific techniques has to be assessed and validated on a multivariate data set.

ACKNOWLEDGEMENT

This research was supported by the German Research Foundation DFG (BRAGECRIM project "Integrating Intelligent Maintenance Systems and Spare Parts Supply Chains (I2MS2C)" – BRAGECRIM 022/2012).

REFERENCES

Angeli, C. (2010). Diagnostic expert systems: from expert's knowledge to real-time systems. In P. S. Sjja & R. Akerkar (Eds.), *Advanced knowledge based systems: Model, Applications & Research* (Vol. 1, pp. 50-73).

Bengtsson, M., Olsson, E., Funk, P., & Jackson, M. (2004). Design of condition based maintenance system - A case study using sound analysis and case-based reasoning. *Condition Based Maintenance Systems - An Investigation of Technical Constituents and Organizational Aspects*. Malardalen University, Eskilstuna, Sweden, 57.

Bonissone, P. P., & Varma, A. (2005). Predicting the best units within a fleet: prognostic capabilities enabled by peer learning, fuzzy similarity, and evolutionary design process. *Fuzzy Systems, 2005. FUZZ'05. The 14th IEEE International Conference on* (pp. 312-318), May 25-25. IEEE.

Cecati, C. (2015). A Survey of Fault Diagnosis and Fault-Tolerant Techniques - Part II: Fault Diagnosis with Knowledge-Based and Hybrid/Active Approaches. *IEEE Transactions On In Transactions Electronics.*

Dai, X., & Gao, Z. (2013). From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis. *Industrial Informatics, IEEE Transactions on, 9(4)*, 2226-2238.

Ghodrati, B. (2005). *Reliability and operating environment based spare parts planning*. Doctoral dissertation. University of Technology, Lulea, Sweden.

Hall, D. L., & Llinas, J. (1997). An introduction to multisensor data fusion*. Proceedings of the IEEE*, 85(1), 6-23.

Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal* Processing, 20.7, 1483-1510.

Li, Q., Gao, Z. B., & Shao, L. Q. (2014). An operating condition classified prognostics approach for Remaining Useful Life estimation. *Prognostics and Health Management (PHM), 2014. IEEE Conference on* (pp. 1-9), June 22-25. IEEE.

Liao, L., & Lee, J. (2009). A novel method for machine performance degradation assessment based on fixed cycle features test. *Journal of Sound and Vibration*, 326(3), 894-908.

Liu, J., Djurdjanovic, D., Ni, J., Casoetto, N., & Lee, J. (2007). Similarity based method for manufacturing process performance prediction and diagnosis. *Computers in industry*, 58(6), 558-566.

Medina-Oliva, G., Voisin, A., Monnin, M., Peysson, F., Léger, J.-B. (2012). Prognostics assessment using fleet-wide ontology. *Annual Conference of the Prognostics and Health Management Society 2012, PHM 2012*, September, p. CDROM.

Mobley, R. Keith (2002). *An introduction to predictive maintenance*. Butterworth-Heinemann.

Monnin, M., Abichou, B., Voisin, A., & Mozzati, C. (2011). Fleet historical cases for predictive maintenance. *The International Conference Surveillance*, 6, 25-26.

Moss, T. R. (1991). Uncertainties in reliability statistics. *Reliability Engineering & System Safety*, *34*(1), 79-90.

Pecht, M., & Jaai, R. (2010). A prognostics and health management roadmap for information and electronics - rich systems. *Microelectronics Reliability*, 50.3, 317-323.

Peng, Y., Dong, M., & Zuo, M. J. (2010). Current status of machine prognostics in condition-based maintenance: a review. *The International Journal of Advanced Manufacturing Technology*, 50.1-4, 297-313.

Rodríguez, G. (2007). Lecture Notes on Generalized Linear Models. URL: http://data.princeton.edu/wws509/notes/

Saxena, A., Sankararaman, S., & Goebel, K. (2014). Performance evaluation for fleet-based and unit-based prognostic methods. *Second European conference of the Prognostics and Health Management society*.

Si, X. S., Wang, W., Hu, C. H., & Zhou, D. H. (2011). Remaining useful life estimation - a review on the statistical data driven approaches. *European Journal of Operational Research*, 213(1), 1-14.

Sikorska, J. Z., Hodkiewicz, M., & Ma, L. (2011). Prognostic modelling options for remaining useful life estimation by industry. *Mechanical Systems and Signal Processing*, *25*(5), 1803-1836.

Tsang, A. H., Yeung, W. K., Jardine, A. K., & Leung, B. P. (2006). Data management for CBM optimization. *Journal of Quality in Maintenance Engineering*, *12*(1), 37-51.

Turrin, S., Subbiah, S., Leone, G., & Cristaldi, L. (2015). An algorithm for data-driven prognostics based on statistical analysis of condition monitoring data on a flee level. *Instrumentation and Measurement Technology Conference (I2MTC), 2015. IEEE International* (pp. 629-634), May 11-14. IEEE.

Voisin, A., Medina-Oliva, G., Monnin, M., Leger, J. B., & Iung, B. (2013). Fleet-wide diagnostic and prognostic assessment. *Annual Conference of the Prognostics and Health Management Society*, October, p. CDROM.

Wang, T., Yu, J., Siegel, D., & Lee, J. (2008). A similarity-based prognostics approach for remaining useful life estimation of engineered systems. *Prognostics and Health Management (PHM), 2008. International Conference* (pp. 1-6), October 06-09. IEEE.