Multi-Sensory Data Analysis and On-Line Evaluation for Advanced Process Control and Yield Optimization in Polymer Film Industry

Multisensorische Datenanalyse und schritthaltende Auswertung für die fortgeschrittene Prozessführung und Ausbeuteoptimierung in der Polymerfilmindustrie

vom

Fachbereich Elektrotechnik und Informationstechnik der Technischen Universität Kaiserslautern zur Verleihung des akademischen Grades

Doktorin oder Doktor der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation

von

Diplom-Wirtschaftsingenieur Michael Kohlert

geb. in Halle/Saale (Deutschland)

D386

Eingereicht am:	06. Mai 2015
Tag der mündlichen Prüfung:	10. Juli 2015
Dekan des Fachbereichs:	Prof. Dr.-Ing. Hans D. Schotten

Promotionskommission

Vorsitzender:	Jun.-Prof. Dr.-Ing. Daniel Görges
Berichterstattende:	Prof. Dr.-Ing. Andreas König
	Prof. Dr.habil.tech. Tamara Chistyakova

Multi-Sensory Data Analysis and On-Line Evaluation
for Advanced Process Control and Yield Optimization
in Polymer Film Industry
Dissertation, TU Kaiserslautern

Abstract

The current procedures for achieving industrial process surveillance, waste reduction, and prognosis of critical process states are still insufficient in some parts of the manufacturing industry. Increasing competitive pressure, falling margins, increasing cost, just-in-time production, environmental protection requirements, and guidelines concerning energy savings pose new challenges to manufacturing companies, from the semiconductor to the pharmaceutical industry.
New, more intelligent technologies adapted to the current technical standards provide companies with improved options to tackle these situations. Here, knowledge-based approaches open up pathways that have not yet been exploited to their full extent. The *Knowledge-Discovery-Process* for knowledge generation describes such a concept. Based on an understanding of the problems arising during production, it derives conclusions from real data, processes these data, transfers them into evaluated models and, by this open-loop approach, reiteratively reflects the results in order to resolve the production problems. Here, the generation of data through control units, their transfer via field bus for storage in database systems, their formatting, and the immediate querying of these data, their analysis and their subsequent presentation with its ensuing benefits play a decisive role.
The aims of this work result from the lack of systematic approaches to the above-mentioned issues, such as process visualization, the generation of recommendations, the prediction of unknown sensor und production states, and statements on energy cost.
Both science and commerce offer mature statistical tools for data preprocessing, analysis and modeling, and for the final reporting step. Since their creation, the insurance business, the world of banking, market analysis, and marketing have been the application fields of these software types; they are now expanding to the production environment. Appropriate modeling can be achieved via specific machine learning procedures, which have been established in various industrial areas, e.g., in process surveillance by optical control systems. Here, *State-of-the-art* classification methods are used, with multiple applications comprising sensor technology, process areas, and production site data. Manufacturing companies now intend to establish a more holistic surveillance of process data, such as, e.g., sensor failures or process deviations, to identify dependencies. The causes of quality problems must be recognized and selected in real time from about 500 attributes of a highly complex production machine. Based on these identified causes,

recommendations for improvement must then be generated for the operator at the machine, in order to enable timely measures to avoid these quality deviations.

Unfortunately, the ability to meet the required increases in efficiency – with simultaneous consumption and waste minimization – still depends on data that are, for the most part, not available. There is an overrepresentation of positive examples whereas the number of definite negative examples is too low. The acquired information can be influenced by sensor drift effects [D$^+$04] and the occurrence of quality degradation may not be adequately recognized. Sensorless diagnostic procedures with dual use of actuators can be of help here. Moreover, in the course of a process, critical states with sometimes unexplained behavior can occur. Also in these cases, deviations could be reduced by early countermeasures.

The generation of data models using appropriate statistical methods is of advantage here. Conventional classification methods sometimes reach their limits. Supervised learning methods are mostly used in areas of high information density with sufficient data available for the classes under examination. However, there is a growing trend (e.g., spam filtering) to apply supervised learning methods to underrepresented classes, the datasets of which are, at best, outliers or not at all existent.

The application field of *One-Class Classification* (OCC) deals with this issue. Standard classification procedures (e.g., k-nearest-neighbor classifier, support vector machines) can be modified in adjustment to such problems. Thereby, a control system is able to classify statements on changing process states or sensor deviations.

The above-described knowledge discovery process was employed in a case study from the polymer film industry, at the Mondi Gronau GmbH, taken as an example, and accomplished by a real-data survey at the production site and subsequent data preprocessing, modeling, evaluation, and deployment as a system for the generation of recommendations. To this end, questions regarding the following topics had to be clarified: data sources, datasets and their formatting, transfer pathways, storage media, query sequences, the employed methods of classification, their adjustment to the problems at hand, evaluation of the results, construction of a dynamic cycle, and the final implementation in the production process, along with its surplus value for the company.

Pivotal options for optimization with respect to ecological and economical aspects can be found here. Capacity for improvement is given in the reduction of energy consumption, CO_2 emissions, and waste at all machines. At this one site, savings of several million

euros per month can be achieved.
One major difficulty so far has been hardly accessible process data which, distributed on various data sources and unconnected, in some areas led to an increased analysis effort and a lack of holistic real-time quality surveillance. Monitoring of specifications and the thus obtained support for the operator at the installation resulted in a clear disadvantage with regard to cost minimization. The data of the case study, captured according to their purposes and in coordination with process experts, amounted to 21,900 process datasets from cast film extrusion during 2 years' time, including sensor data from dosing facilities and 300 site-specific energy datasets from the years 2002–2014.
In the following, the investigation sequence is displayed:

1. In the first step, industrial approaches according to Industrie 4.0 and related to Big Data were investigated. The applied statistical software suites and their functions were compared with a focus on real-time data acquisition from database systems, different data formats, their sensor locations at the machines, and the data processing part. The linkage of datasets from various data sources for, e.g., labeling and downstream exploration according to the knowledge discovery process is of high importance for polymer manufacturing applications.
2. In the second step, the aims were defined according to the industrial requirements, i.e. the critical production problem called "cut-off" as the main selection, and with regard to their investigation with machine learning methods. Therefore, a system architecture corresponding to the polymer industry was developed, containing the following processing steps: data acquisition, monitoring & recommendation, and self-configuration.
3. The novel sensor datasets, with 160–2,500 real and synthetic attributes, were acquired within 1-min intervals via PLC and field bus from an Oracle database. The 160 features were reduced to 6 dimensions with feature reduction methods. Due to underrepresentation of the critical class, the learning approaches had to be modified and optimized for one-class classification, which achieved 99% accuracy after training, testing and evaluation with real datasets.
4. In the next step, the 6-dimensional dataset was scaled into lower 1-, 2-, or 3-dimensional space with classical and non-classical mapping approaches for downstream visualization. The mapped view was separated into zones of normal and abnormal process conditions by threshold setting.

5. Afterwards, the boundary zone was investigated and an approach for trajectory extraction consisting of condition points in sequence was developed, to optimize the prediction behavior of the model. The extracted trajectories were trained, tested and evaluated by *State-of-the-art* classification methods, achieving a 99% recognition ratio.
6. In the last step, the best methods and processing parts were converted into a specifically developed domain-specific graphical user interface for real-time visualization of process condition changes. The requirements of such an interface were discussed with the operators with regard to intuitive handling, interactive visualization and recommendations (as e.g., messaging and traffic lights), and implemented.

The software prototype was tested at a laboratory machine. Correct recognition of abnormal process problems was achieved at a 90% ratio. The software was afterwards transferred to a group of on-line production machines.

As demonstrated, the monthly amount of waste arising at machine M150 could be decreased from 20.96% to 12.44% during the application time. The frequency of occurrence of the specific problem was reduced by 30% related to monthly savings of 50,000 EUR.

In the approach pertaining to the energy prognosis of load profiles, monthly energy data from 2002 to 2014 (about 36 trajectories with three to eight real parameters each) were used as the basis, analyzed and modeled systematically. The prognosis quality increased with approaching target date. Thereby, the site-specific load profile for 2014 could be predicted with an accuracy of 99%. The achievement of sustained cost reductions of several 100,000 euros, combined with additional savings of EUR 2.8 million, could be demonstrated.

The process improvements achieved while pursuing scientific targets could be successfully and permanently integrated at the case study plant. The increase in methodical and experimental knowledge was reflected by first economical results and could be verified numerically. The expectations of the company were more than fulfilled and further developments based on the new findings were initiated. Among the new finding are the transfer of the scientific findings onto more machines and even the initiation of further studies expanding into the diagnostics area.

Considering the size of the enterprise, future enhanced success should also be possible for other locations. In the course of the grid charge exemption according to EEG, the energy savings at further German locations can amount to 4–11% on a monetary basis and at least 5% based on energy. Up to 10% of materials and cost can be saved with regard to waste reduction related to specific problems. According to projections, material savings of 5–10 t per month and time savings of up to 50 person-hours are achievable. Important synergy effects can be created by the knowledge transfer.

Kurzfassung

Heutige Verfahren der industriellen Prozessüberwachung, Ausschussreduktion und Prognose von kritischen Verfahrenszuständen sind in Teilen der produzierenden Industrie bislang unzureichend gelöst. Der steigende Wettbewerbsdruck, sinkende Margen, Kostenwachstum, Flexibilisierung, *Just-in-Time*-Fertigung, Umweltschutzauflagen und nachhaltige Energieeinsparrichtlinien stellen Produktionsunternehmen von der Halbleiter- bis zur pharmazeutischen Industrie vor neue Herausforderungen.
Neue intelligentere Technologien, angepasst an den heutigen Stand der Technik, eröffnen Unternehmen verbesserte Möglichkeiten diesen Gegebenheiten entgegenzuwirken. Wissensbasierte Ansätze eröffnen dazu Wege, die bislang nicht in vollem Umfang genutzt wurden. Der *Knowledge-Discovery*-Prozess zur Wissensgenerierung beschreibt ein solches Konzept. Aus dem Verständnis für die Produktionsprobleme leitet es anhand realer Daten Erkenntnisse ab, bereitet die Daten auf, überführt diese in evaluierte Modelle und reflektiert in einem *Open-Loop*-Ansatz wiederkehrend Ergebnisse, um Produktionsprobleme zu lösen.
Dabei spielen die Erhebung der Daten mittels Steuerungen, ihre Übertragung über Feldbus zur Speicherung in Datenbanksystemen, ihre Formatierung, die zeitnahe Abfrage dieser Daten, ihre Auswertung sowie die unterstützende Darstellung und ein daraus abgeleiteter Nutzen eine entscheidende Rolle.
Die Ziele dieser Arbeit ergeben sich aus der fehlenden methodischen Herangehensweise an die genannten Problemstellungen wie die Prozessvisualisierung, das Vorschlagswesen, die Vorhersage von unbekannten sensorischen und Produktionszuständen und Aussagen über die Energiekosten.
Aus dem kommerziellen und dem Forschungsbereich werden ausgereifte statistische Werkzeuge für die Datenvorverarbeitung, -analyse, -modellierung und die Berichterstattung angeboten. Der Einsatzbereich dieser Softwaretypen erstreckt sich seit Beginn vom Versicherungsbereich, dem Bankenwesen und der Marktanalyse bis hin zum Marketing und weitet sich heutzutage verstärkt in Richtung des Produktionsumfeldes aus. Die passende Modellierung kann mittels spezifischer maschineller Lernverfahren erfolgen, die sich in verschiedensten Bereichen der Industrie, z.B. in der Produktionsüberwachung mittels optischer Kontrollsysteme etabliert haben. Dabei kommen *State-of-the-art*-Klassifikationsverfahren zum Einsatz, deren Anwendungsfelder von der Sensorik über Prozessfelder bis hin zu Standortdaten vielfältig sind.

Produktionsunternehmen wollen Prozessdaten verstärkt ganzheitlich überwachen, wie z.B. Sensorausfälle oder Prozessabweichungen, und Abhängigkeiten identifizieren. Ursachen für Qualitätsprobleme müssen in Echtzeit aus etwa 500 Eigenschaften einer hochkomplexen Herstellungsmaschine erkannt und selektiert werden. Daraus abgeleitet müssen zusätzlich Verbesserungsvorschläge für den Mitarbeiter an der Maschine generiert werden, um Qualitätsabweichungen rechtzeitig entgegenzusteuern.
Die Deckung dieses Bedarfs an Effizienzsteigerung – bei gleichzeitiger Verbrauchs- und Ausschussminimierung – ist jedoch abhängig von Daten, die größtenteils nicht vorliegen. Es besteht eine Überrepräsentation von Gutbeispielen, die Anzahl tatsächlicher Fehlerbeispiele ist dagegen zu gering. Akquirierte Informationen können durch sensorische Drifteffekte [D+04] beeinflusst werden. Eine Degradation wird möglicherweise nicht ausreichend erkannt. Hier können sensorlose Diagnoseverfahren unter Doppelverwendung von Aktoren Hilfestellung bieten. Auch Prozessverläufe können kritische Zustände enthalten, deren Verhalten bisweilen ungeklärt ist. Auch hier können Abweichungen durch frühzeitige Gegenmaßnahmen reduziert werden. Dazu ist der Aufbau von Datenmodellen unter Verwendung geeigneter statistischer Methoden von Vorteil.
Bisweilen stoßen herkömmliche Klassifikationsmethoden an ihre Grenzen. Überwachte Lernmethoden werden meistens auf Gebieten hoher Informationsdichte eingesetzt, wo genügend Daten für die betrachteten Klassen vorhanden sind. Es entwickelt sich jedoch ein Trend (z.B. *Spam Filtering*) hin zur Methodenanwendung auf unterbesetzte Klassen, deren Datensätze allenfalls als Ausreißer oder gar nicht vorhanden sind.
Das Anwendungsfeld der Ein-Klassen-Klassifikation (OCC) beschäftigt sich mit dieser Fragestellung. Standardklassifikationsverfahren (z.B. k-nearest-neighbor classifier, support vector machine) können mittels Modifikation an solche Problemstellungen angepasst werden. Dadurch ist ein Überwachungssystem in der Lage, Aussagen über sich verändernde Prozesszustände oder sensorische Abweichungen zu klassifizieren.
Der oben beschriebene *Knowledge-Discovery*-Prozess wurde am Fallbeispiel der Kunststofffolienindustrie in einer für diese neuartigen Studie bei der Mondi Gronau GmbH praktiziert und mittels reeller Datenerhebung, Datenvorverarbeitung, Modellierung, Evaluierung und Deployment in Form eines Vorschlagswesens umgesetzt. Dazu waren die Fragen nach den Datenquellen, den Datensätzen, deren Formatierung, den Übertragungswegen, den Speichermedien, der Abfragesequenz, den angewendeten Methoden der Klassifizierung, deren Anpassung an die Problemstellungen, der Evaluierung

der Ergebnisse, der Konstruktion eines dynamischen Kreislaufs und der finalen Implementierung im Produktionsprozess mit Mehrwert für das Unternehmen zu klären.
Ein bisheriges Defizit waren schwer zugängliche Prozessdaten, die auf verschiedenen Datenquellen verteilt und nicht verknüpft in einigen Bereichen zu erhöhtem Analyseaufwand und fehlender ganzheitlicher Echtzeit-Qualitätsüberwachung führten. Das Monitoring von Spezifikationen und daraus abgeleitete Hilfestellungen für den Werker an der Anlage bildeten hier einen klaren Nachteil im Umgang mit der Kostenminimierung. Die nach ihren Zielsetzungen in Abstimmung mit den Verfahrensexperten erfassten Daten der Fallstudie belaufen sich auf 21.900 Prozessdatensätze der *Cast*folienextrusion aus einem Zeitraum von 2 Jahren, darunter sensorische Daten von Dosiereinrichtungen und 300 Standortenergiedaten aus den Jahren 2002–2014.
Im Folgenden ist der Ablauf dargestellt.

1. Im ersten Schritt wurden industrielle Anwendungen aus dem Resort Industrie 4.0 und Big Data betrachtet hinsichtlich des Einsatzes von statistischen Software Suites und deren methodischen Funktionalitäten. Der Fokus lag dabei auf dem Framework der Echzeit-Datenakquisition aus Datenbanksystemen, der unterschiedlichen Datenformatierung, deren Zuordnung zu Maschinenkomponenten und Weiterverarbeitung. Die Verknüpfung der Daten spielt dabei eine entscheidende Rolle, um Prozesseigenschaften mit Label zu kennzeichnen und deren Exploration im Sinne der Wissensgenerierung nutzbar zu machen für ein neues Anwendungsfeld im Bereich der Polymerindustrie.
2. Im zweiten Folgeschritt wurden auf Grund der industriellen Anfordungen mit Ausrichtung auf ein spezielles Materialproblem die Zielstellung definiert, die Vorhersage von kritischen "Abrissen" unter Verwendung von maschinellen Lernverfahren. Dazu musste eine System-Architektur entwickelt werden, die den gesamten Prozessablauf von der Datenakquisition, über das Monitoring & Vorschlagswesen, bis hin zum Selbst-Konfigurierenden System in Gänze darstellt. Die Architektur wurde an die industriellen Anforderungen der Polymerindustrie ausgerichtet und schrittweise untersucht.
3. Im Zuge der Untersuchung wurden unbehandelte sensorische Prozessdaten, 160–2.500 mögliche reale und synthetische Attribute, in 1-Minuten-Intervallen über SoftSPS und Feldbus aus einer Oracle-Datenbank ausgelesen und nach Merkmalsreduktion von 160 Dimensionen auf 6 Dimensionen in maschinelle Lernmodelle überführt. Auf Grund der Unterrepräsentation von kritischen Mengen des Fallbeispiels, wurden die Verfahren

modifiziert auf Ein-Klassen Klassifikation ausgerichtet und optimiert. Die Verfahren wurden mit realen Daten antrainiert, getestet und evaluiert mit einer Genauigkeit von 99%.
4. Im nächsten Schritt wurden unterschiedliche klassische und nicht-klassische Skalierungsverfahren untersucht und anhand eines Rekonstruktionsfehlers bewertet, um den 6 dimensionalen Ergebnisraum auf 1, 2 oder 3 Dimensionen zu reduzieren und nachfolgend zu visualisieren. Das Ergebnis-Mapping wurde folgend mittels Limitierung in unterschiedliche Zonen unterteilt, die den normalen vom kritischen Prozessdatenraum trennen.
5. Innerhalb einer weiterverfolgten 2 dimensional betrachteten Grenzzone konnten anschließend durch ein eigens entwickeltes Extraktionsverfahren Trajektorien, Abfolgen von nachgelagerten Zuständen, bestimmt werden zur Verbesserung der zeitlichen Vorhersage. Die extrahierten Trajektorien wurden mittels maschinellen Lernverfahren in Modelle überführt, antrainiert, getestet und evaluiert bei einer Genauigkeit von 99%.
6. Der letzte Schritt überführt die vorher gewonnenen Verfahren und Entwicklungen in das Anwendungsfeld, ein eigens entwickeltes Software-Interface zur Echtzeit-Visualisierung von Prozessveränderungen. Dazu wurden die Anforderungen an solch ein Interface am Fallbeispiel der Polymerindustrie aufgenommen und implementiert. Die Funktionen fokussierten sich auf intuitive Bedienbarkeit, interaktive Visualisierung und professionelles Vorschlagswesen mittels Nachrichtendienst und Ampelsystem.
Darauf aufbauend wurde der Software Prototyp an einer Labormaschine getestet und die korrekte Erkennung von kritischen Prozesszuständen mit einer Genauigkeit von 90% geprüft und nachfolgend auf eine Gruppe von On-line Produktionsmaschinen überspielt.

Nachweislich konnte der monatliche Ausschuss der Prototypmaschine im Zeitraum der Anwendung von 20,96% auf 12,44% gesenkt werden. Das spezifische Problem wurde in der Häufigkeit seines Auftretens um etwa 30% reduziert, was einer monatlichen Einsparung von 50.000 EUR entspricht.
Beim Ansatz der energetischen Vorhersage von Lastgangverhalten wurden monatliche Energiedaten von 2002 bis 2014 (etwa 36 Trajektorien mit jeweils drei bis acht realen Merkmalen) zugrunde gelegt, ausgewertet und methodisch modelliert. Die Prognosegüte stieg mit zunehmender Annäherung an das Zieldatum. Dabei konnte das Standort-Lastgangverhalten mit einer Genauigkeit von 99,9% für 2014 vorhergesagt werden. Eine

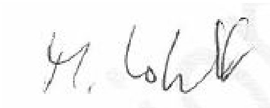
nachhaltige Kostenersparnis von mehreren 100.000 EUR und eine daran gekoppelte zusätzliche Einsparung von 2,8 Mio. EUR traten nachweislich ein.

Die aus wissenschaftlichen Zielen erreichten Prozessverbesserungen konnten dauerhaft erfolgreich am Fallbeispiel integriert werden. Die methodischen und experimentellen Erkenntnisse konnten in ersten wirtschaftlichen Ergebnissen reflektiert und anhand von Zahlen belegt werden. Die Erwartungen des Unternehmens wurden dabei mehr als übertroffen und daran anknüpfende weiterführende Entwicklungen wurden gestartet, darunter die Übertragung der wissenschaftlichen Erkenntnisse auf weitere Maschinen und darüber hinausgehende Untersuchungen im Bereich der Diagnostik. Mit Blick auf die Größe des Unternehmens lassen sich hier zukünftige Erfolge auch auf andere Standorte übertragen. Die Energieeinsparung kann an weiteren deutschen Standorten im Zuge der Netzentgeltbefreiung nach EEG monetär zwischen 4 und 11% sowie energetisch mindestens 5% betragen. Bei der Ausschussoptimierung können für spezifische Problemfälle bis zu 10% an Materialien und Kosten eingespart werden. Nach Hochrechnungen sind hier 5–10 t an Materialersparnis pro Monat sowie bis zu 50 Personalstunden an Zeitersparnis erreichbar. Mit der Überführung können weitreichende Synergieeffekte geschaffen werden.

Declaration

I declare that the following dissertation is the product of my own work. It has not been submitted before for any examination or work in any other institution, and all the sources I have used or quoted have been indicated and acknowledged as complete references.

Kaiserslautern, 23.09.2015
Michael Kohlert



Acknowledgements

Deeply gratefully acknowledgement to Prof. Dr.-Ing. Andreas König, holder of the Chair for Integrated Sensor Systems (Department of Electrical Engineering and Computer Engineering at the Technical University of Kaiserslautern), for suggesting the general topic of this research work, advise and support, discussions and continued encouragement. His recommendations, considerations, and critical reviews mainly contributed to the success of this work.

I would like to express my great acknowledge to Prof. Dr.habil.tech. Tamara Chistyakova, holder of the Chair for Computer Aided Design and Control (St. Petersburg State Institute of Technology/ Technical University), for her interest in my research and for her reviewing efforts value adding to this work.

For the "Führungskreis 1" at the Mondi Gronau GmbH (2014) obtains my special acknowledgement for unlimited support.

I would like to thank my colleagues from Production, Engineering, IT and Administration at the Mondi Gronau GmbH for their outstanding encouragement and for sharing their fundamental knowledge in the course of these projects.

Last but not least I want to thank my parents and my brother for their patience and additional ideas for continuing progress. Their contributed experience intensively supported my research effort.

Table of Contents

List of Figures

List of Tables

List of Algorithms

1 Introduction

In 2012, the German chemical industry represented about 20% of the European chemical industry, with Bayer and BASF holding the highest market values in 2014. The European chemical industry is divided into the areas of consumer chemicals (12.3%), petrochemicals (24.7%), specialties (25.3%), basic organics (13.4%), and polymers (24.3%). The corresponding sector breakdown by Cefic Chemdata International from 2012 is displayed in Fig. 1 [HP13].

Petrochemicals, specialties, and polymers represent the highest market share in Europe, with 74.3%. The number of foreign direct investment projects in Europe is highest for software and business services. The Experton Group[1] predicts a 54% investment growth in information technologies (IT) for Germany in 2015 [Zil14]. Gartner[2] estimates 4.4 million new IT jobs in big data challenges for 2015, called "... new opportunities for: transforming decision-making; discovering new insights; optimizing the business; and innovating their industries" [Pet12].

In German production plants, the production department (52%) is currently more busy with "Industrie 4.0" than the management (45%) and IT departments (34%) [[Zil14] [ZW14] [Zac12] [Stu11]], which should be balanced at least.

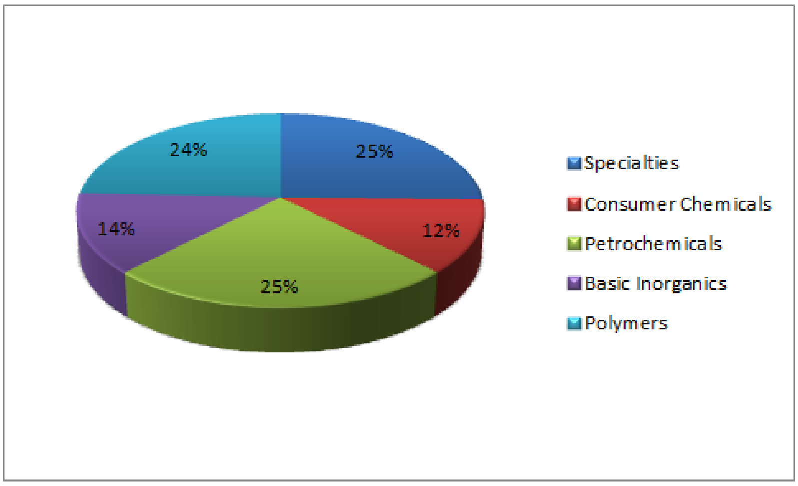


Figure 1 According to Cefic: EU Chemicals Industry Sales by Sectoral Breakdown (2012) [HP13]

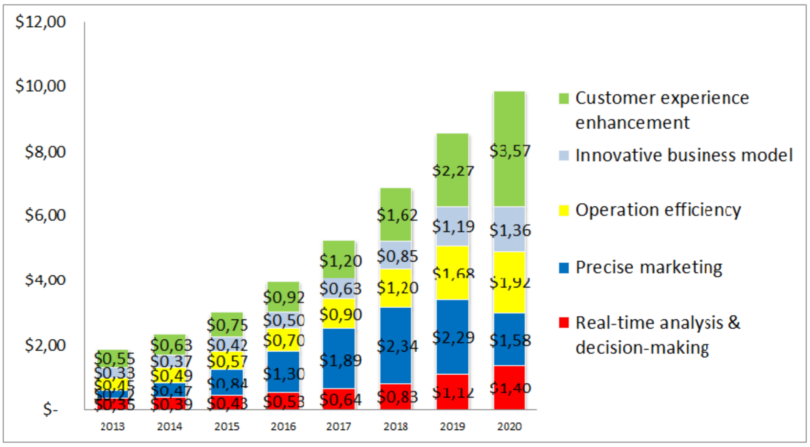


Figure 2 According to Heavy Reading: Big Data Analytics Market Size by Business Category [in Billions $] [Ban13]

According to Heavy Reading (2013) the demand for "Real-time analysis & decision-

[1] Experton, Provider of market research studies
[2] Gartner, Provider of market research studies

making" will raise by 200% from 2014 to 2020 (Fig. 2). Today's manufacturing industry recognizes the potential of analytical research for their production lines to raise productivity. In case of the Fraunhofer Institute, the industrial growth depends on its high production share since the depression in 2008/2009, which helped Germany to recover faster than other countries [G+13]. The Organization for Economic Co-operation and Development (OECD) has declared that a stronger focus on innovations is needed (investment into Research & Development (abbr.: R&D)), which would lead to higher productivity and competition (high association between R&D and productivity). The governmental involvement as described by the OECD (2014) should help in supporting competitiveness in different ways (to generate opportunities) in order to be effective for diverse industrial areas. For example, lowering the energy consumption provides a higher productivity, which is typically true for industries where new technologies and processes are promoted [OEC14].
McKinsey (2014) reports that innovations "will become the basis of competition and growth for individual companies, enhancing productivity [...]" [M+14]. The industrial image increasingly changes to smaller production batches, higher flexibility, non-centralized automation, and mass customization.
The motivations for improvement are widespread, e.g., the optimization of slow, interfering intra-company workflows for faster execution of business processes to gain more time for important tasks, sustainable reduction of the energy consumption of obsolete machine equipment plotted against the return on investment (ROI), or media acquisition in plants with high energy consumption to benefit from the Renewable Energy Law (transl.: EEG [Bun14]), or the net stability[3] in Germany, realizing several millions of energy cost per annum. Furthermore, sequence planning optimization at production lines to reduce material waste, setup times and cost, which today is made manually and not holistically, constitutes about 5–20% of the margin (more than 200k EUR per machine/ annum). The predictable periodicity of maintenance for downtimes (unexpected occurrence), machine component malfunctions (not monitored) and processing problems (unknown machine condition changes) to avoid waste decreases yields losses of more than 5% (savings of several millions EUR, kg material, and kWh energy in plants of more than 800 employees are possible per month). The application of newer technologies for faster and more precise client acquisition may raise the efficiency by 40% [SAS14], but may also take a lot of time and investment today for small success in the future. These fields are chosen as typical examples with high demand for yield optimization necessary for many plants.
The German government proclaimed a new digitalized high-technology strategy ('Industrie 4.0') carried out by cyber-physical production systems (CPPS), which are autonomous systems with self-optimization (self-x) and smart technologies for all industrial types, to ensure a head start for the German industry in the world. Typically, high amounts of data are used for yield optimization in manufacturing industries by optimizing processing routines for faster reaction when unknown states occur, or for a more efficient reduction of energy consumption at obsolete non-environmental production plants [[Win13] [Bun15]].
Picking out one of the above-mentioned improvement fields, manufacturing plants have high demand for the prediction of machine component malfunctions and processing problems, as discussed in the following Chapters. Typical manufacturing plant problems

[3] Bundesnetzagentur (2013)

occur because of missing real-time monitoring, insufficient boundary settings, or lacking dependency recognition of sufficient sensor data points within production lines due to inadequate technical experience of the operators.

1.1 Survey on industrial Big Data Methods and Developments

Today's advanced analytical methods in the manufacturing industry are mainly focused on quality systems for in-line (e.g., defects or thickness) and laboratory (e.g., strength ratio) control measurements. Basically, the supporting operators' tasks aim is satisfying the clients' specifications and ordinal standards with multi-class methodology.
Besides quality and recipe, the processing data constitutes the main part within production, contributing datasets of 30–2000 possible parameters from a production line (e.g., a calendering line), depending on the machine complexity. Occasionally, about 5% of these parameters are acquired and monitored, disregarding the rest. On the other hand, unknown machine states or process behaviors are not predictable.
In addition, *State-of-the-art* data analysis methods are not fully satisfactory, with problems arising from abnormal datasets caused by sensor errors (e.g., missing data) and unknown process conditions. Optimized boundary specifications are not established for classes with low amounts of data.
As disclosed by expert knowledge, erroneous analytical results from inefficiently linked data sources and machine parts lead to flawed conclusions regarding the causes of insufficient quality. Thus, wrong machine settings, repeatedly inadequate interpretation of quality changes, recurrent abnormal conditions, sensor property deviations, limited real-time evaluations, unfiltered information types, and non-existing operator recommendations basically lower the yield and increase capital expenditures due to higher material waste, energy consumption and labor cost, thus decreasing productivity.
For this reason, the *State-of-the-art* classification methods have to be modified according to the issues specified above, to achieve the industrial requirements. Basically, the occurrence of unknown objects has to be predicted in time, sensor problems need to be discovered in real time, the measurement accuracy needs to achieve the specifications, and an open-loop control is to be implemented to guarantee the reproducibility of the approach for practical use.

1.2 Goals of the Thesis

The major goal of this thesis is the practical investigation of big data methods and research work on dedicated methods from information processing and computational intelligence, e.g., classifiers for novelty and anomaly detection implemented within predictive process control systems, to advance manufacturing processes of polymerfilm industry, and to better meet aggravating industrial requirements and challenges.
A particular focus will be on learning and adapting work from microelectronics manufacturing and related yield optimization for this aim [[M+01], [KG05],[BMB14]].
Sustainable manufacturing processes have a high demand for surveillance of novel conditions and sensor deviations. Incidental fabrication states, abnormal sensor data divergence, and unknown machine condition lead to widely differing, randomized product quality results. A schematic sensory location overview is given in Fig. 3, displaying typical sensor location points for machine data extraction at a polymer film line.

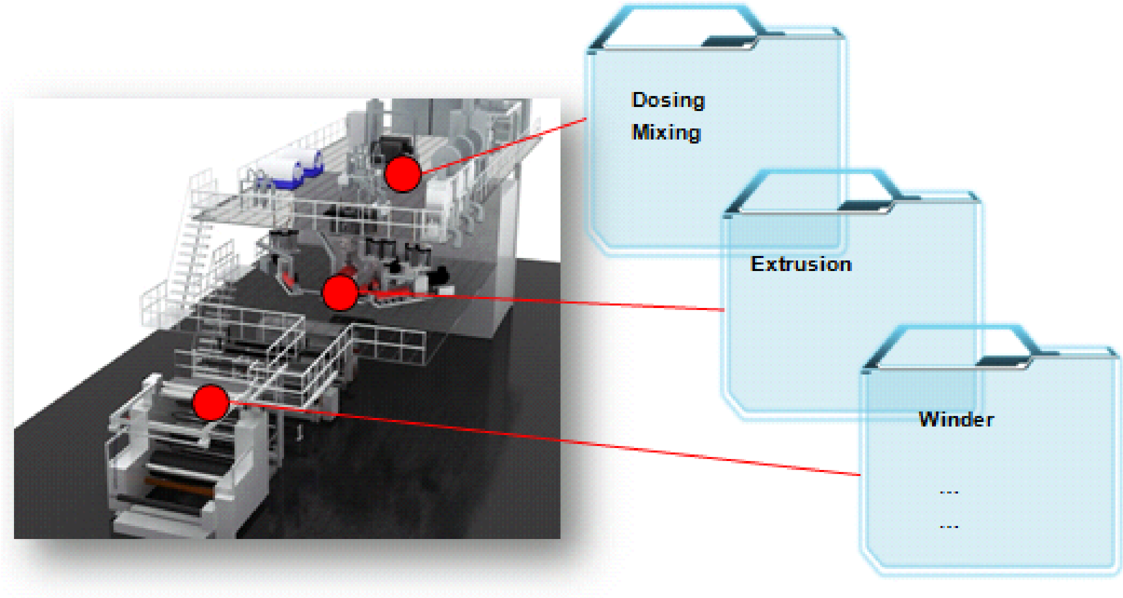


Figure 3 Schematic sensory location points at polymer film line collectable via programmable logic controller and data collector

The usual implementation will take a programmable logic controller (PLC), which sends sensor-derived datasets through a data collector to a database (e.g., Oracle); the database will be read out by analytical programs via a query code for further feature screening. For newer machines, the sensor-derived datasets will be already readably integrated into the machine control, directly at the machine interface (e.g., Data-Gate). The process condition changes and sensor deviations, e.g., drift effects, resulting from environmental influences will lead to material waste and yield loss. For process monitoring improvement to archieve early reaction, novel investigations as, e.g., trajectory approaches shall be examined in this work, optimizing anomaly detection as a rapid alert system. This shall maintain a stable construction against environmental changes in real time. The development of a representative case study, potentially reusable and extendable for related manufacturers in the field, as a research vehicle for real-time computer-aided condition monitoring shall be mainly focused.
Software classification methods, adjustment to real sensor properties, process condition monitoring, and open-loop recommendations, adopted from semiconductor industry, deployed in a practical use case, will be regarded in this work strongly corresponding to Industrie 4.0 approaches, e.g., computational intelligence, smart factory, self properties (Self-X), Cyber-Physical-Production-Systems (CPPS), and human-machine-interaction for raising productivity, efficiency and the automation level.

1.3 Thesis Structure

In Chapter 2, the *Knowledge-Discovery-Process* is defined and a general overview is given on data analysis applications, functionality, data acquisition, and surveillance types. General buzz words like *Big Data*, *Self-X*, *Smart Technologies* and *Cyber-Physical Production Systems* are newly defined within, in this case, the government strategy of *'Industrie 4.0'*.
Furthermore, *State-of-the-art* process controls in the manufacturing industries and

the acquired datasets are explained, followed by a practical approach in the polymer film industry described in Chapter 3, focusing on the manufacturing process, data acquisition, and typically occurring problems.
In Chapter 4, a particular, novel approach for anomaly and novelty detection in the field of sensor properties with *One-Class Classification* (OCC) is investigated, starting with an overview of the field of applications with regard to data acquisition and processing tasks and different adapted methods. Based on the previous chapter, the trajectory process visualization in Chapter 5 analyzes the process visualization support through trajectory behavior in industrial applications of high-dimensional space.
A system architecture for the complete methodolodical enrivonment is developed in Chapter 6, afterwards an overview of the conducted experiments and the achieved results are presented in Chapter 7. In Chapters 8 and 9, the results are discussed and a laboratory data acquisition *con*dition real-time *mon*itoring system (*ConMon*) is realized for off-line data acquisition, processing, and investigation and is afterwards implemented online into the production process for process control with interactive visualization and a suggestion system. Chapter 10 concludes with novel contributions.

2 Knowledge-Discovery-Process in Manufacturing Industry

The automated *Knowledge-Discovery-Process* in the manufacturing industries provides crucial advantages compared to the conventional analysis of production data. The main focus lies on the near real-time prediction of process states for different production machines and the recognition of sensor property deviations. The early identification of failures helps to prevent material waste and leads to yield optimization [[PJ05] [Per07] [Dan05] [BG05] [Bha05]].

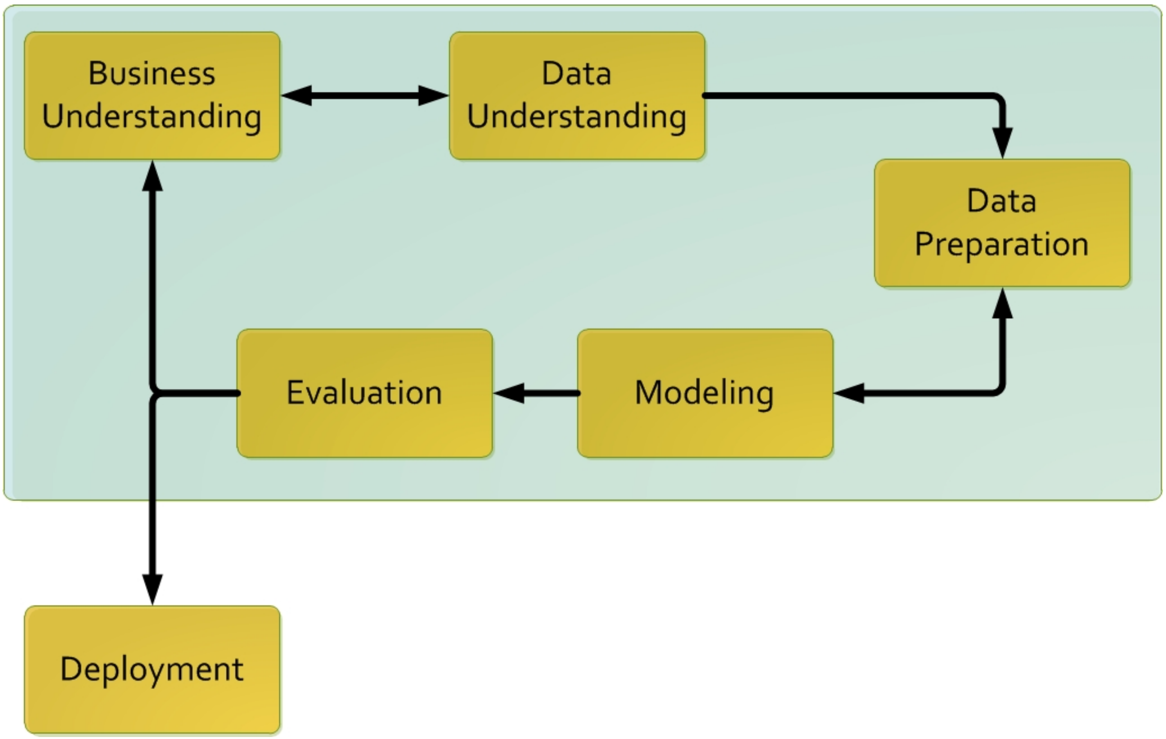


Figure 4 According to CRISP cycle: data understanding, preparation, modeling, evaluation and deployment [$C^{+}07$]

The field of *Knowledge-Discovery-Processes* (KDP)[4], as illustrated in Fig. 4 has evolved due to the emergence of huge datasets which cannot be easily achieved from fast results. A *Knowledge-Discovery-Process* describes the scientific model of problem solving in six steps: (1) understanding the business, (2) understanding the data, (3) data preparation, (4) modeling (data mining), (5) evaluation, and (6) deployment. In more detail, knowledge discovery describes the search, preparation, extraction, modification, cleaning, analysis, reporting, visualization, and interpretation of datasets. The knowledge discovery process (Fayyad, 1996) with focus on databases also includes the data mining step [[Fay96] [Mar12]]. Data mining is defined as part of knowledge discovery for the identification and extraction of unknown, non-trivial, unpredicted, and important information from huge datasets. It is a bottom-up approach using different methods for pattern recognition. Essentially, it is the contribution of data from heterogeneous data sources and a range of methods to the classification, approximation, prediction, association,

[4] CRISP-DM (Cross-Industry Standard Process for Data-Mining - www.crisp-dm.org)

clustering and description of models, with classification, approximation and prediction pertaining to modeling and feature selection, and association, clustering and description to the relation between variables. Thereafter, the development of a data mining model can be realized with reverse feedback from the process for an update function, which also allows the prediction of upcoming process steps. An open-loop approach provides a continuous surveillance of currently generated knowledge [[Mar08] [Wan07] [Twe14]]. The design of open-loop control systems within the manufacturing industry includes parts like a programmable controller, a data collector, the database, and software for querying, transformation and analyzing, adapted to the process environment. The process framework resulting from the product instructions and the process settings established to achieve certain quality characteristics define the particular specifications. These requirements are part of the six basics in the CRISP cycle in Fig. 4.

The on-line and off-line discovery processes follow the same sequence and adaptations to feature changes, if necessary. The accuracy depends on the trained recognition of process states; the recognition rate should be at least 90%, which is the minimum recognition ratio for optical defects measurement systems in polymer film plants [OCS15]. The scanning iteration depends on the production process, which is different for different working stations. The recognition rate needs to remain constant for different products over time, and the safety codes must prevent calculation failures, to ensure that the operators' confidence in the monitoring system does not diminish. The operator must be able to examine the important status data in real time, to obtain an impression of the occurring process deviations. Each product needs to be adapted to the complete control system (Fig. 5), requiring an ordinary adjustment. In addition, the conversion from the laboratory to the online monitoring system is based on standardized communication protocols (e.g., TCP/IP) and strictly prohibits any influence on the primary productive systems, to avoid conflicts.

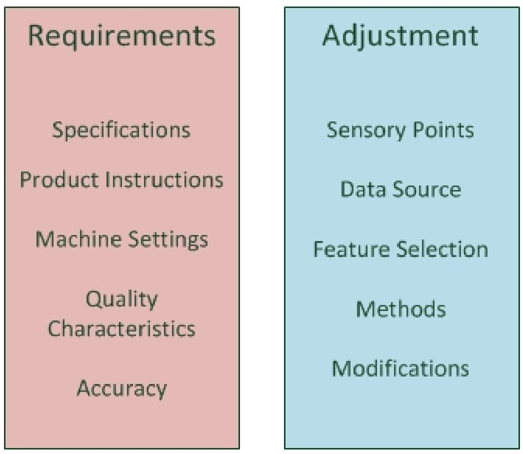


Figure 5 Open-Loop framework for the process monitoring cycle

Connecting all the above mentioned components within a loop generates a processing cycle for an open-loop monitoring system for manufacturing surveillance.

The following paragraphs of this Chapter consider the conceptual parts of the processing cycle, regarding the joint technology roadmap of the NAMUR and VDE/ VDI on process sensors and Mayato and Rexer on software information and methods [[Nec10] [Rex13] [V$^+$09] [Gne14]].

2.1 Data Acquisition and Adaptation

Data acquisition in manufacturing plants starts at sensor points converting analog to digital signals for monitoring, storage, or post-analysis purposes. Missing or faulty data caused by transmission and sensor problems from environmental influences lead to wrong conclusions and recommendations, with expensive consequences.

Data sources, integrating technologies, administration efforts, quality requirements, and dataset sizes have increased since 2011. The demand for analytical databases collecting manufacturing data via a programmable logic controller (PLC) grows, for faster process understanding and pattern recognition in higher data volumes. Concatenation of different format types, drill-down functions, aggregation and database accession methods through querying routines in real time or, overall, the combination of sensor devices, networks, databases and processing applications are summarized by the *Internet of Things* (IoT) technologies.
The following Subsection deals with the different data acquisition parts and their interaction for processing cycles in industrial approaches, from the PLC to the database, the data formats, and the accessing routines [[Mar12] [Mar11] [H$^+$14] [VS14] [BtH07] [Gla14]].

2.1.1 Programmable Logic Controller and Data Collector

Today's manufacturing machines are controlled by computer systems using computerized numerical control (CNC) for programming sequences and PLC for code sequence operations and machine control, visualized via a communication command interface. Known types are, e.g., S7 or S5, installed depending on the functionality of the machine and the quantity of sensor locations. A typcial machine "VLAN" consists of, e.g., HP, Insys or MOXA switches, RJ45 or LWL connection between the switches and machine terminal interfaces and PLC, assigned by a network address, a subnet mask, a standard gateway, a broadcast and an IP range. The PLC processing data can be visualized via a software interface from the vendor (e.g., Datagate). Connected to a data collector (PLC) via ethernet TCP/IP, the whole setup works like a cache: In case of connection interruptions, the datasets are not lost (PLC storage) but, upon reconnection, can be stored in a database system in table format (e.g., Oracle) [Ora14b].

2.1.2 Database Sources

Databases (constructed in C++, Java, PHP, PEARL) are distinguished as single-table, relational, object-oriented, network, document, and hierarchical types, with subclasses. For the selection of specified data fields, tables within sources are necessary, preferably from storage and in real time. The overall implementation of data warehouse solutions became popular for the prevention of the risk of failures occurring when collecting data from productive systems. Holistic data source implementations (e.g., data warehouse) are rare in small manufacturing plants; for these, the established, normally non-linked systems (e.g., MES, SAP, Quality Sources, Energy-Management) offer non-stacked inhomogeneous datasets for analysis, making cause search difficult.

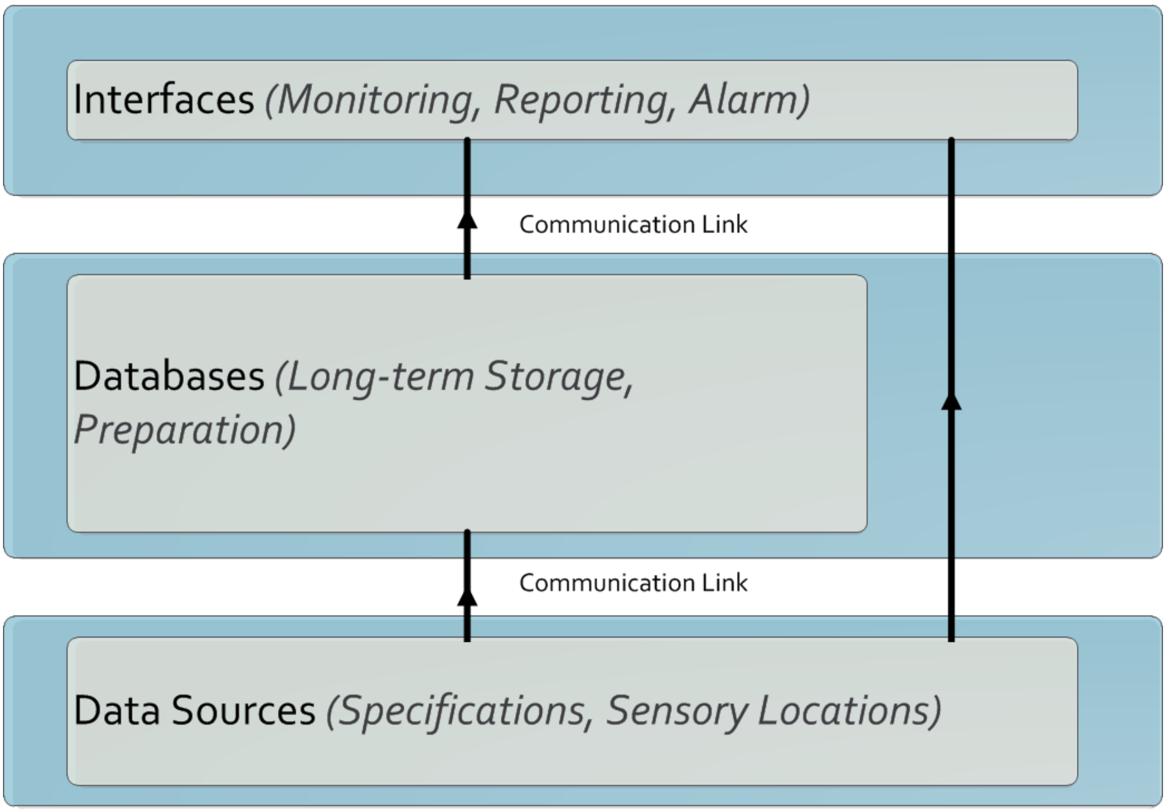


Figure 6 Exemplary data transfer overview within manufacturing plant (data sources, databases, interfaces)

Figure 6 shows an exemplary overview of a data transfer structure within a polymer manufacturing plant including the PLC in the production machines, a data collection point, an Oracle database, and visualization tools. The PLC at the bottom provides data to a data collector, which sends the data to an Oracle database, from where it passes to the manufacturing execution system and to different visualization and reporting tools [Ora14a].
Surveillance of sensor data or detection of sensor property faults within the production process in real time is not granted in manufacturing plants. The sensor points need to be surveyed to prevent damage or waste at the machines. The diagnostics installation is not *State-of-the-art*, due to lack of money and experience.

2.1.3 Data Types and Formats

Data formats describe the type of data stored in a database, accessed or read out for further processing. Queried datasets from Oracle databases are readable in character, numeric, DATE, LOB, RAW, LONG RAW, ROWID, and UROWID, depending on their informational transaction reason. Constraints on data sources determine changes in the data types and the set relations between tables. [Ora14b]

2.1.4 Accessing Databases

The retrieval of datasets from data sources like Oracle is performed with query languages. Different types of query languages exist, e.g., Structured Query Language (SQL) or Extensible Markup Language (XML). SQL is mostly used to access relational databases; in contrast, XML is used for websites, MDX (Multidimensional Expression) for OLAP databases, and DMX (Data Mining Extensions) for additional standards. The selection of datasets from Oracle tables within a relational database via the command sql is shown in the following sample:

$$'SELECT * FROM(SELECT fa, rollennr FROM data.t - object)' \tag{1}$$

By inserting specific names and command variables into the displayed string, a query is started on tables from the sources for collecting data, which can be deposited in different tables, concatenated to one table, calculated or monitored [Sta14c].

2.2 Data-Mining Software Suites

Software tools for data mining are available as suites or commodities, providing specialism versus standardization. The interpretation of results is done by process experts and statisticians [[Kus06] [Kue99] [H$^+$05] [H$^+$01] [Dav14] [Bra02]]. According to Mayato (2010) about 150 data mining applications exist that are in use for explorative fields. As presented in Fig. 7, the Rexer study from 2011 shows a rated selection of the most popular tools on the market.

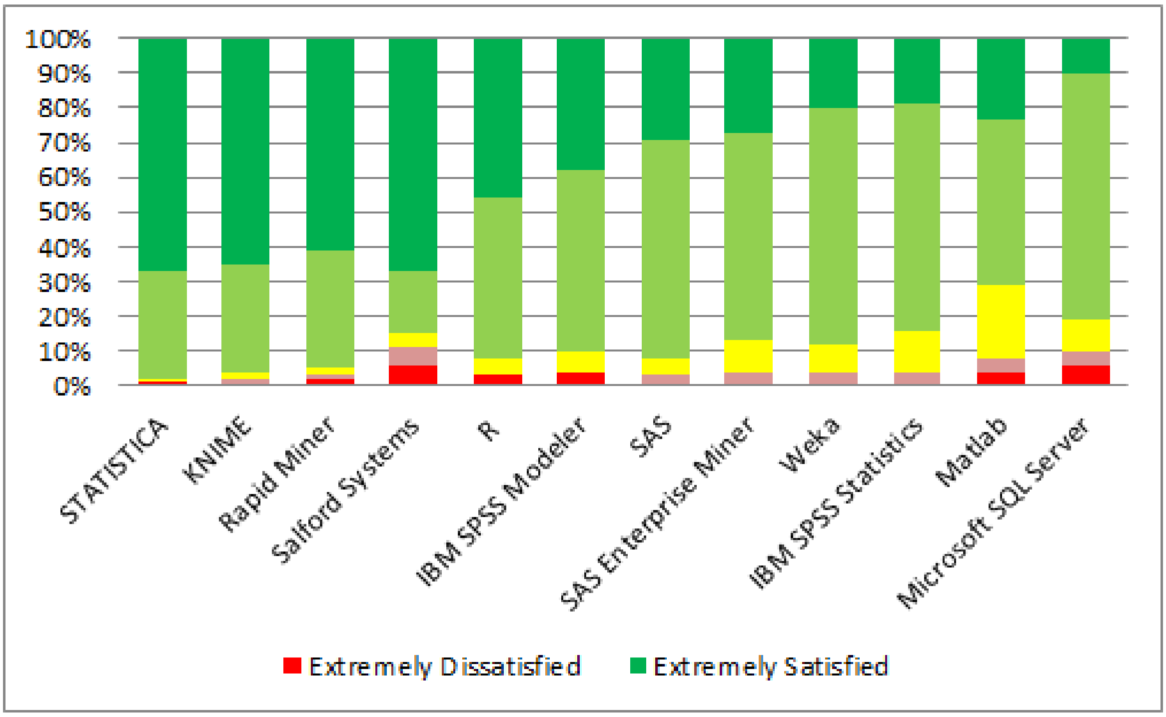


Figure 7 According to the Rexer Study from 2011 - Different software tools and their customer satisfaction ratio

According to this, STATISTICA, KNIME, Rapid Miner, Salford Systems, and R lead the operator satisfaction rating. The study is based on reports on analytical tools by 1300 users from 60 countries. STATISTICA (StatSoft) was voted as the most popular commercial tool in 2012, in the 13th yearly KDNuggets Software Poll [KDN12], therefore examined for further analysis in this work.
Data mining suites can be differentiated into commercial ones like SAS Enterprise Miner, IBM SPSS Modeler, and STATISTICA Data Miner and open-source ones like Rapidminer, KNIME, and WEKA. KXEN Analytics Framework is the leading suite in

automation, ranked as self-acting tool with high solution platform. Classical instruments are QuickCog, Viscovery SOMine, Prudsys Realtime Decisioning Engine, Bissantz Delta Master and, from SAP BW 7, Data Mining Workbench, Oracle 11g Data Mining, and Microsoft SQL Server 2008 Analysis Service, as shown in Fig. 8 from the study of Mayato (2010).

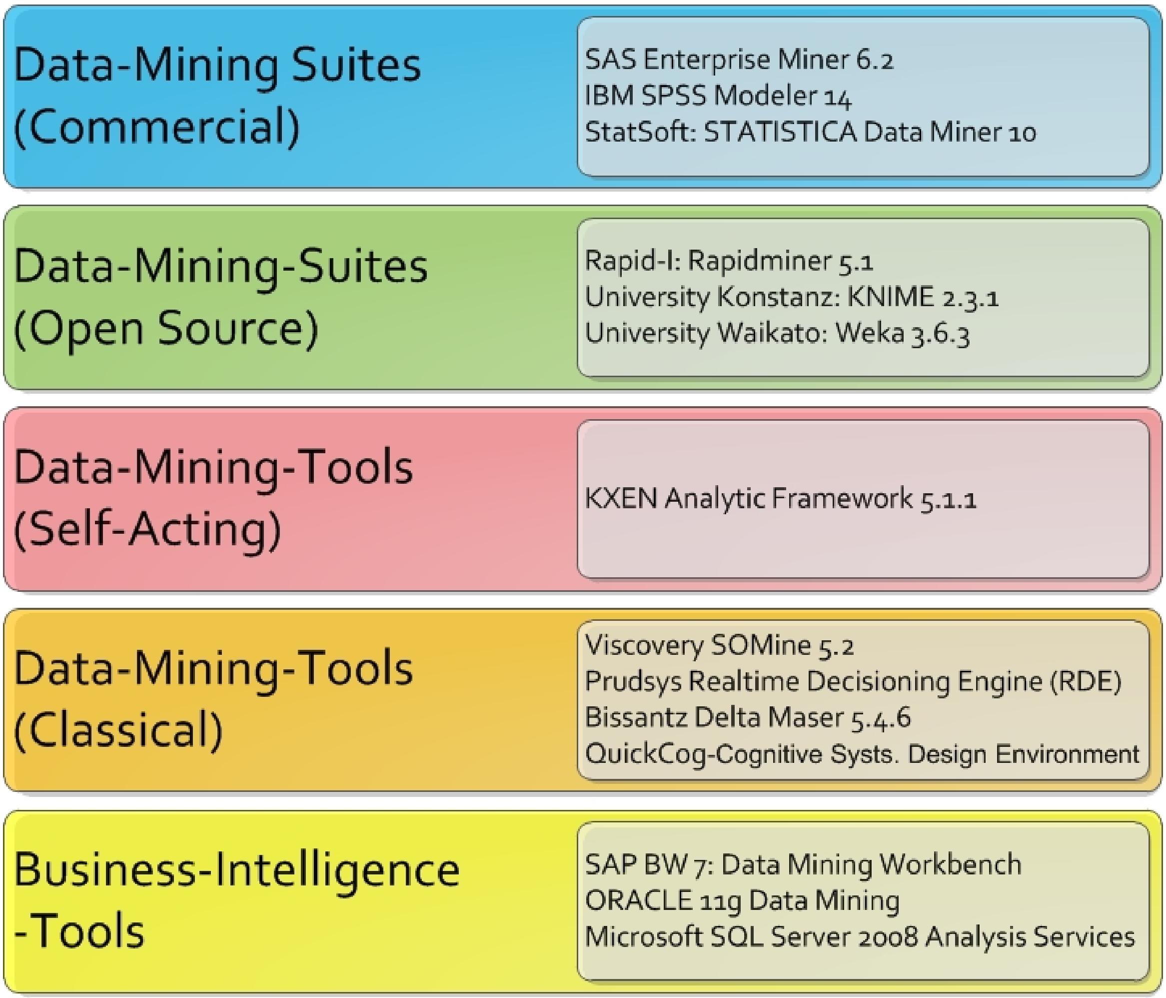


Figure 8 According to the Mayato Survey 2010: Modified comparison of analytical tools

Regarding the analytical capabilities, each tool provides a range of methods matching the desired *Knowledge-Discovery* model; for example, the QuickMine Toolbox in QuickCog is particularly useful for sophisticated visual analytics [K$^+$98].
The Rexer Analytics Data Miner Survey 2008 identified stability, system performance with big datasets, and quality of results as the most important ones of the “Top Priorities for Software Selection”. The Aberdeen study included intuitive handling as most significant feature [[Nec10] [Rex13] [Abe12]].
The following Subsections give short summaries of the deployed analytical tools, such as STATISTICA by StatSoft GmbH, SAS Visual Analytics by SAS Institute, Inc., Matlab & Simulink by MathWorks, Inc., Rapidminer by Rapidminer, and QuickCog by A. König et al. These tools were chosen regarding their industrial and research benefit published in several papers and at conferences worldwide. Their field of applications is wide spread from semiconductor to pharmaceutical industry. Regarding Table 1, STATISTICA, SAS Visual Analytics, and Rapidminer are mainly applied in commercial

approaches, whereas Matlab & Simulink, and QuickCog are strengthend in research & development fields. For selection purpose the software features visualization, *One-Class* Methods, Workspace, and Interface Connection were focused.

Table 1 Compared software types by properties, as visualization, one-class methods, solution help, and configuration help

Criteria: / Software:	Visualization	One-Class Methods	Solution Help	Configuration Help
STATISTICA	Graphs, SPC, more	No	Support	Setting Options Shown
SAS Visual Analytics	Graphs, SPC, more	No	Support	Setting Options Shown
Rapidminer	Graphs	No	Sample Applications	Setting Options Shown
Matlab & Simulink	Graphs, SPC, more	Add-on (Tax, Duin)	Sample Applications	Documentation Needed
QuickCog	Graphs	Integrated (NOVAS, NOVCLASS)	Sample Applications	Setting Options Shown

The intuitive handling was left out as subjective property. Finally QuickCog, and Matlab & Simulink were chosen due to their *One-Class* ability, STATISTICA due to their visualization, and configuration environment.
In the following Subsections the previous compared software types are explained.

2.2.1 STATISTICA

STATISTICA by StatSoft GmbH[5] is a statistical software with a modular structure and with a huge selection of methods for data analysis. From basic, advanced and industrial statistics to clustering and classification methods, text mining, rule extraction, graphical visualization, in e.g., statistical process control charts, and final deployment, STATISTICA offers an extensive analytical variety of ad hoc or workspace model generation options for experienced users. The connection module to external data sources through querying routines provides a simple monitoring assistance for data acquisition from, e.g., manufacturing controls or manufacturing execution systems.

2.2.2 SAS Analytics

With 36.2%, SAS[6] was the largest market shareholder for advanced analytics in 2012 [Zac12]. SAS Visual Analytics offers a wide range of statistical methods and visualization tools for huge data analysis and reporting. The software suite allows ad hoc processing and modeling within workspace environments and further code compiling for

[5] StatSoft GmbH, http://www.statsoft.com (Accessed: 14.01.2015)
[6] SAS Institute, Inc., http://www.sas.com (Accessed: 14.01.2015)

database integration of analytical sequences. Data acquisition from several data sources, concatenation for processing, modeling and implementation within an open-loop affords all elements of knowledge discovery. [Ste09]

2.2.3 Matlab & Simulink

MathWorks, Inc.[7], offers an interpreter programming software for technical calculations, simulations and dynamical design recommended for experienced users (e.g., in research). Different toolboxes provide a high selection of various codes and methods for, e.g., building up data acquisition cycles, trained models, or control charts, and fast deployment into a complete standalone solution.

2.2.4 Rapidminer

Rapidminer[8] is an analytical software for data mining, machine learning, and predictive analytics. The workspace environment offers a variety of methods for statistical analysis, to be used in processing routines.

2.2.5 QuickCog

QuickCog[9] is a system design tool for automated visual inspection projects and for general pattern recognition applications with adaptation possibilities to tasks with little or no expert knowledge. The modular environment offers *State-of-the-art* classification methods as well as feature space visualization and interactive analysis. The key features can be assigned on a workspace structure for modeling purposes of high-dimensional datasets supported by application samples.

2.3 Implementation into a Dynamic Open-Loop System

Starting with sensor data transmission through PLC to data collectors, additional analog sensor points converted to digital ones can be added consecutively, stored in an Oracle database (necessary for machines with more than 50 sensory locations) selected for processing via the sql query.
The standard *Knowledge-Discovery-Process* modified for the manufacturing industry can be described in a complete process diagram (Fig. 9), according to Fig. 4 mentioned at the beginning of this Chapter.

[7] Mathworks, Inc., http://www.mathworks.com (Accessed: 14.01.2015) [The15]
[8] Rapidminer, https://rapidminer.com (Accessed: 14.01.2015)
[9] QuickCog, http://www.quickcog.de (Accessed: 14.01.2015)

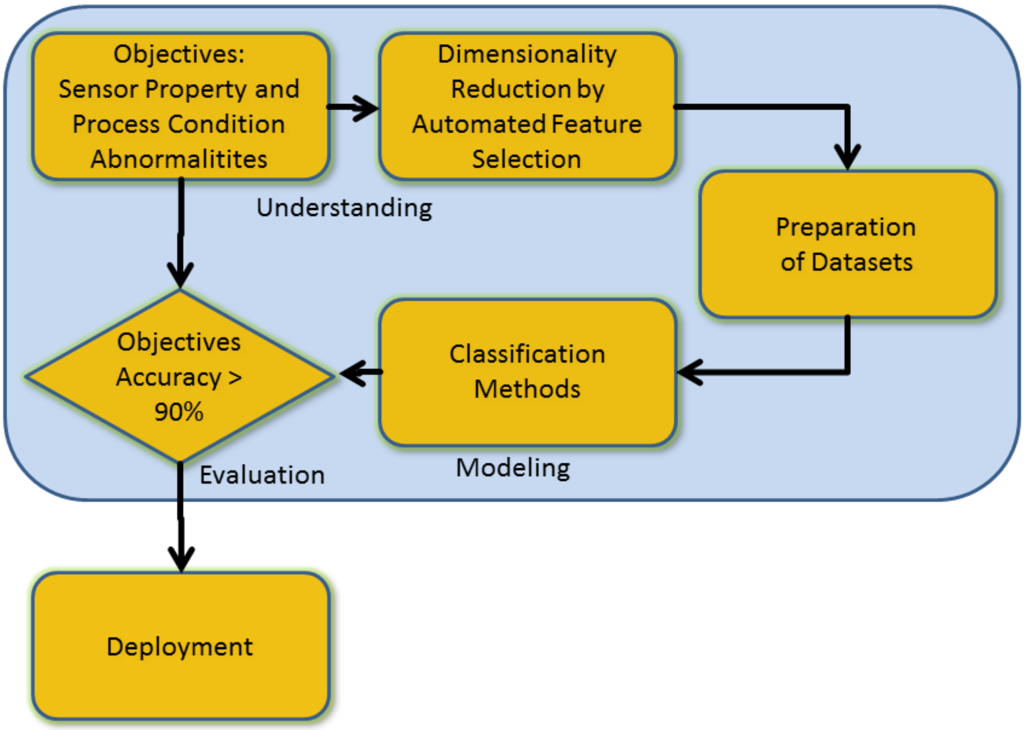


Figure 9 According to *CRISP* cycle the *Knowledge-Discovery-Process* in a manufacturing approach for sensor property and quality prediction [C+07]

The objectives, as sensor properties and process condition abnormality shall be defined in the beginning of the cycle; a feature selection for dimensionality reduction shall be made in the following step. Training, validation, and testing subsets will be prepared for modeling with classification methods, evaluated afterwards and deployed to standalone applications if the accuracy compared to other common measurement systems accuracy complies with the objectives' requirements. The features dedicated to pattern recognition shall be examined. Novel approaches on supervised OCC types will be investigated to be explained in the following Chapters. Such modified classification methods are approved and do not cause any problems in classes with low amounts of data. Thus, abnormality detection can be improved with OCC methods for the detection of process deviations.

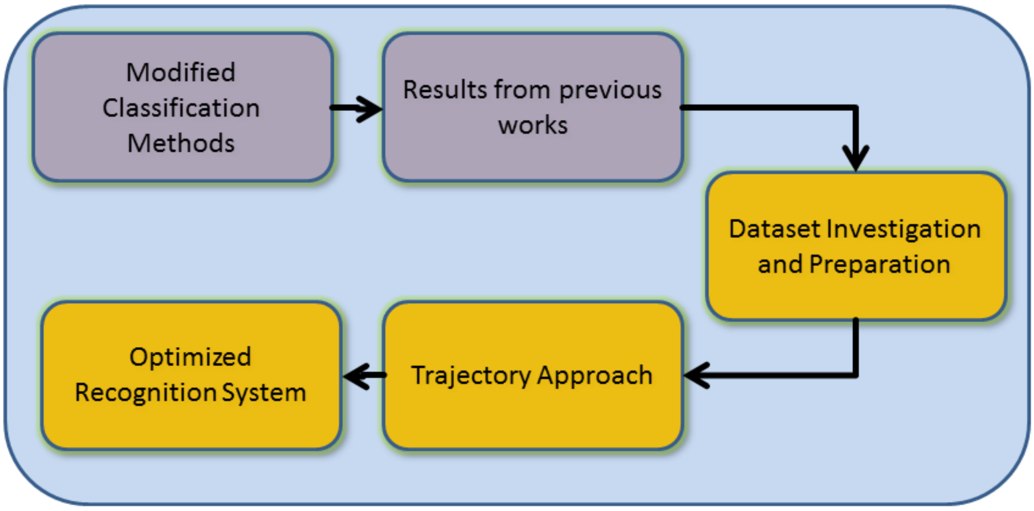


Figure 10 Outlier- and trajectory classification in a two-stage classification approach

Based on the optimized classification method and the open-loop setup, the first stage of recognition shall shift the object into the outlier or the normal class. Here, this stage is called outlier classification.
To separate the normal trend conditions from faulty trends, a second classification shall be carried out on the normal-condition datasets, denoted here as trajectory approach, shown in Fig. 10. The repetition frequency of this loop depends on the manufacturing process (e.g., 2 -h of production per product, time interval 1 min).

2.4 Discussion

Data analysis is divided into the descriptive, the explorative, and the confirmative field. The confirmative field uses a given search space, the explorative field, an open search space. Data mining aims at uncovering unknown relations with automated sequences. The present-day tools offer statistical ratios beside the exploration of datasets, data preparation, and other functions. Data extraction from different sources and transformation to a common format is the first step, followed by selection, exploration, modification, analysis, and interpretation of the results. Through exploration, a basic understanding of patterns can be achieved. The modification step adjusts data and missing data and recognizes redundancy. The analysis step is supported by a range of methods, applications, and model ratings.
The search for hidden patterns in big datasets and feature extraction (correlation) to make conclusions on new datasets are the basic steps for predictive modeling. The degree of success depends on the data quality, as inefficient preparation of datasets leads to erroneous examination results. Problems caused by low data quality, e.g., when data is missing or faulty, lead to underdetermined classes for comparison of different operation conditions. Besides, changes in sensor location may lead to analog signals from sensor points influencing the accuracy of the results. This can lead to analysis failure and wrong conclusions. Currently, the data quality in manufacturing industries is not surveyed in closed or open loops [Nec10].
The software functionality has increased during the last years, due to higher performance

for huge datasets and the adaptation of *State-of-the-art* methods to specific classification problems such as *One-Class* issues.

3 A Practical Approach in the Polymer Film Industry

With a market share of 24.3%, the polymer industry constitutes one of the main chemical industry sectors worldwide, beside the petrochemical sector with 24.7% [HP13] [Sta14b] [Sta14a]. Products made of plastic play a role in everybody's life, e.g., in cars, packaging, pharmaceutical products, or in the semiconductor industry. Therefore, this work focuses on real production data from polymer plants to investigate a real-life application.
The following Chapters start with a description of the polymer film production process, first giving a holistic overview of the general processes and then moving on to the materials level for insights on polymer materials with regard to their application fields and demand developments. Afterwards, the production machines, from calendering to extrusion, are described with a focus on the aggregates level, providing an impression of their control within electronic data-processing circuits. The next Chapters deal with process data division, sensor location examination, quality data, and process specifications, concluding in challenges for a research vehicle company from the polymer industry, to achieve waste reduction as well as trajectory and energy prediction.

3.1 Polymer Film Production Process

The polymer film production process is divided into the different types of materials, the production machines, the specification settings including customer requirements, and the environmental influences. Although the processes differ for different polymers, their converting steps are similar, as a polymer material is heated up and cooled down, transformed into a solid mass and prepared for further tasks. These sequences are repeatedly executed, with potentially increasing complexity; see Fig. 11, displaying an exemplary three-stage polymer process with compounding, extrusion, and lamination stages for elastomer production, resulting in 50-μm films for the hygienics industry.

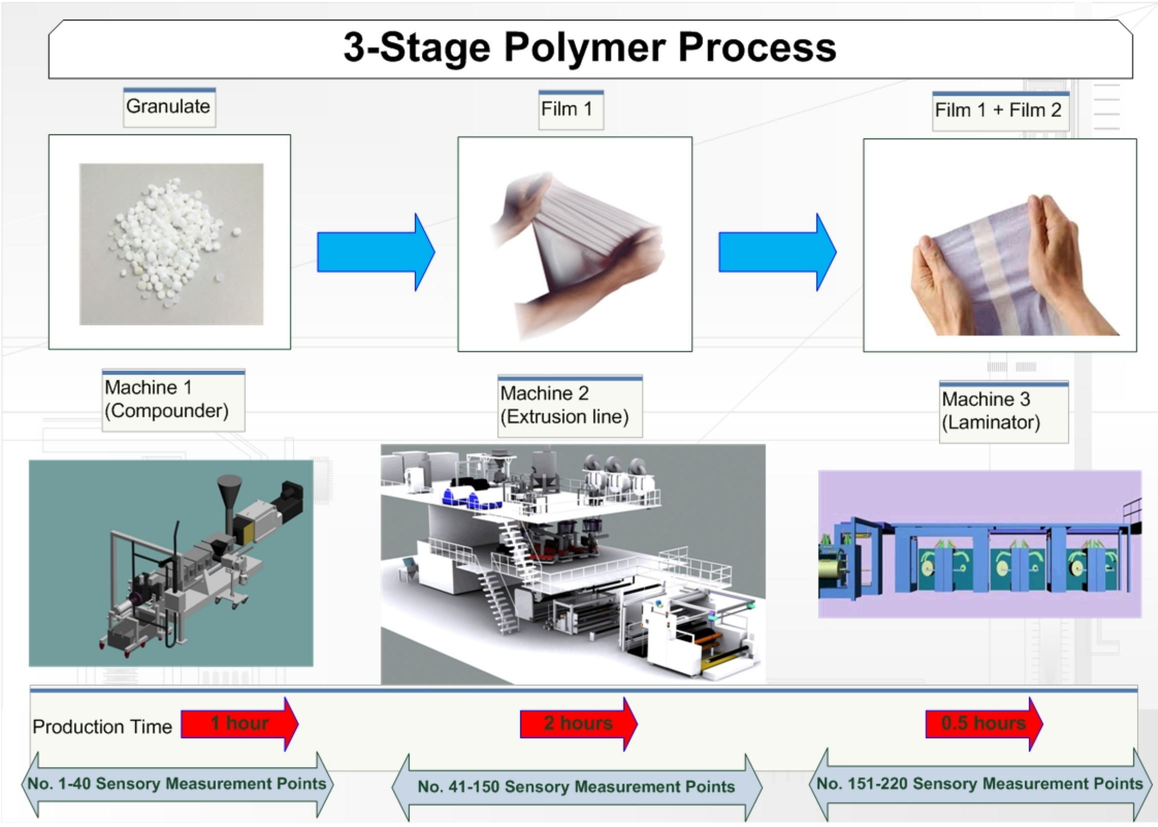


Figure 11 3-stage polymer film process for elastomer products including production time and sensory points [KK12]

As presented in Fig. 11, the polymer granulate is melted up by a compounder at 180°C, then filtered, washed and dosed into an extruder, melted up again at 190°C, converted to a film by a nozzle, rolled up, and laminated with nonwoven in the next converter stage. This case shows the typical parts of the complex machines used within the polymer industry to produce different products. The different types of polymer materials used in the various polymer machines are explained, giving a short overview of the electronic programmable logic controller (PLC) processing.

3.1.1 Polymer Materials (SBS, PE, PP, PET, PVC)

The European plastics demand in 2012 is displayed in Fig. 12, by segment and resin type. With 39.9%, ‘Packaging’ has the biggest market share, followed by ‘Others’ with 26.6% and ‘Building & Construction’ with 20.3%. The main demand types are PE (Polyethylene), PP (Polypropylene), and PET (Polyethylene terephthalate) for ‘Packaging’, PVC (Polyvinyl chloride) for ‘Building & Construction’, and PP and PUR (Polyurethane) for ‘Others’[10][11].

[10] Index of Abbreviations
[11] SBS = Styrol-blockcopolymere

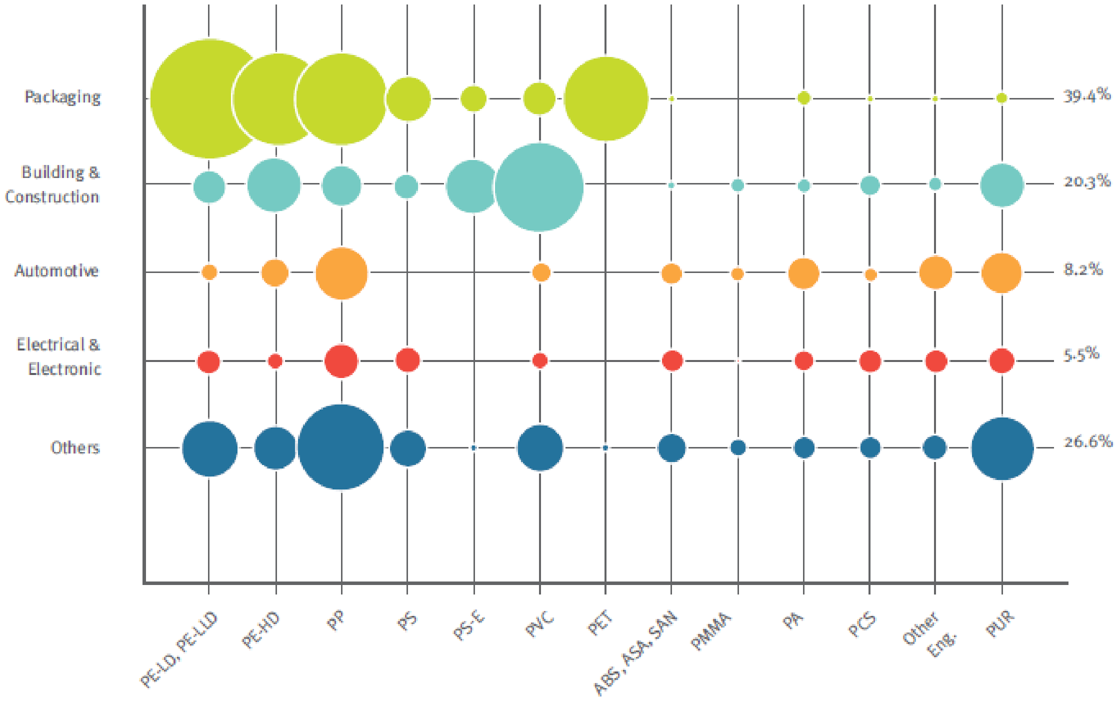


Figure 12 According to: European plastics demand* by segment and resin type 2012, source: PlasticsEurope (PEMRG) / Consultic / ECEBD (* EU-27+N/CH)

Plastics recycling and energy recovery reached 61.9% in 2012. Landfill disposal decreases every year by about 5.5%, in relation to 45.9 Mt of plastics produced per year in Europe [Pla13]. The different types of polymer materials are displayed in Fig. 13 according to Sagel [Sag12], showing the predominantly manufactured plastics, PP, PVC, and PE subtypes (HDPE - Polyethylene, high density; LLDPE - Polyethylene, linear low density; LDPE - Polyethylene, low density), and those produced in lesser amounts: PET, PS (Polystyrene), ABS (Acrylonitrile butadiene styrene), and PC (Polycarbonate). PE (more precisely its subtypes) was the most used polymer in 2012 worldwide, covering 37% of the world polymer demand of 211 million metric tons.

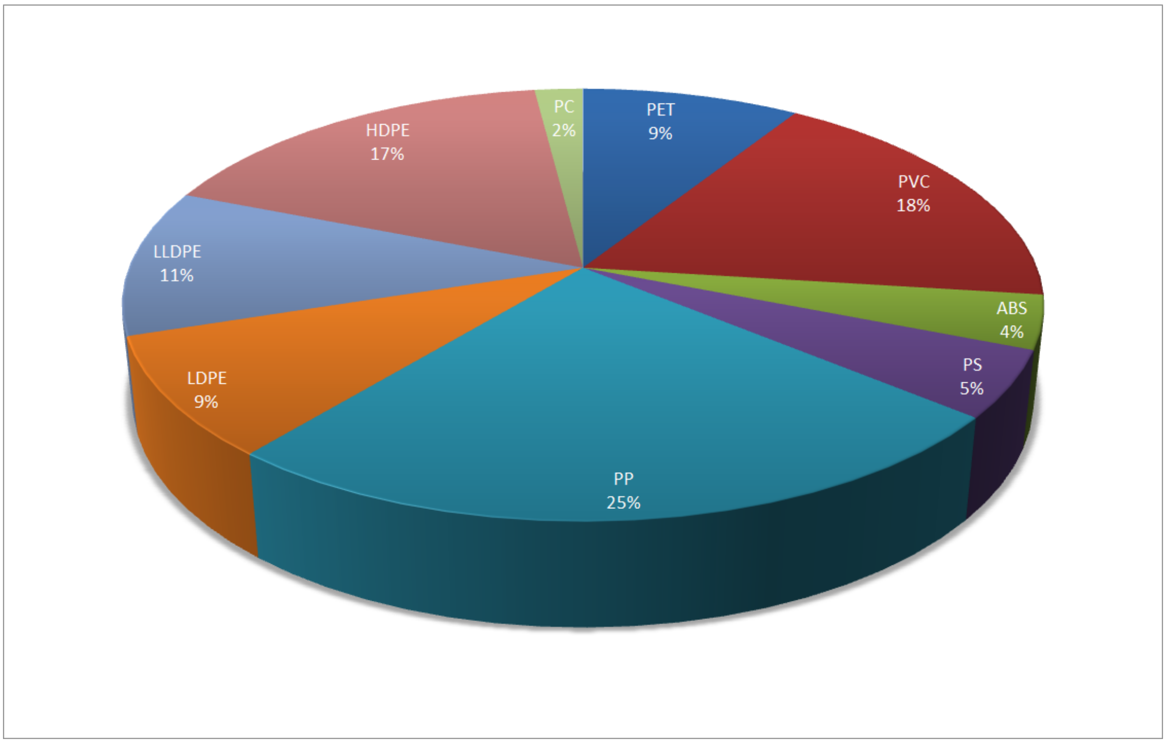


Figure 13 According to Sagel: Polyethylene Global Overview 2012 [Sag12]

According to Plastics Europe[12] (2013), PP is used for, e.g., bumpers and folders, PVC for boots and windows, PE subtypes for containers, caps, bags and cables, PET for bottles, PS for glass frames and cups, and ABS for bricks, as displayed in Fig. 14.

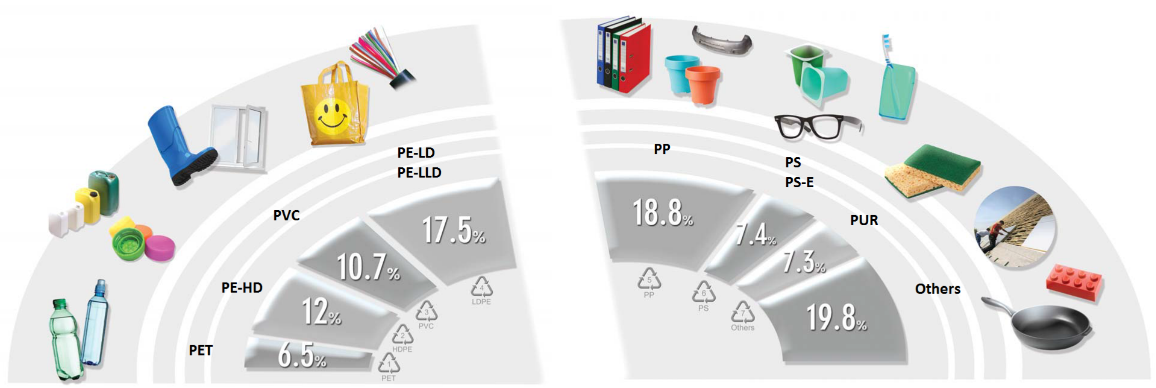


Figure 14 According to: European plastics demand* by segment and resin type 2012, source: PlasticsEurope (PEMRG) / Consultic / ECEBD (* EU-27+N/CH) [Pla13]

Their properties make them suitable for different kinds of applications, in various combinations. For specific research on polymers the book "Handbuch Urformen" by Spur et al. (2014), and "Kunststoffe" by Domininghaus (2012) is recommendable [[S+14] [D+12]].

[12] European plastics demand (EU27+N/CH) by resin type 2012 [Pla13]

3.1.2 Production Machines (Compounder, Extruder, Calender)

The typical processes in the polymer film industry are injection, rotational molding, calendering, cast-extrusion, blown film extrusion, foaming, and additional process segments like compounding and lamination. Due to the different process types, various machines, usually named like the process, and procedures exist for producing polymer products. Some typical processes are described in this Subsection.

Compounder

In addition to the polymer processes below, the compounding process related to the extrusion process melts up dosed polymers within an extruder and forms smaller granulates from the raw material (Fig. 15).

Figure 15 Animated compounder construction with machine interface in front

This process usually does not produce films and is located upstream of the extrusion processes, as a preprocessing stage for material size reduction and washing [S+14].

Injection Moulding

The injection molding process produces parts by injection of different polymer materials into a mold. It is the most common method in the polymer industry and allows fast production of a multitude of products. Up to two PLCs control such machines, depending on their complexity, and these are usually operated as a group, by one operator [S+14].

Cast-Extrusion

Similar to the calendering process, the cast extrusion process starts with dosing stages for heated mixing of granular polymers (e.g., SBS = Styrol-blockcopolymere, PP), feeding the extruders (e.g., 3, 5, 7 extruders, each one for different film layers (60–220°C)). The formed film coming out of a wide (1–4 m) nozzle is rolled up and slitted at the winder, as shown in Fig. 16.

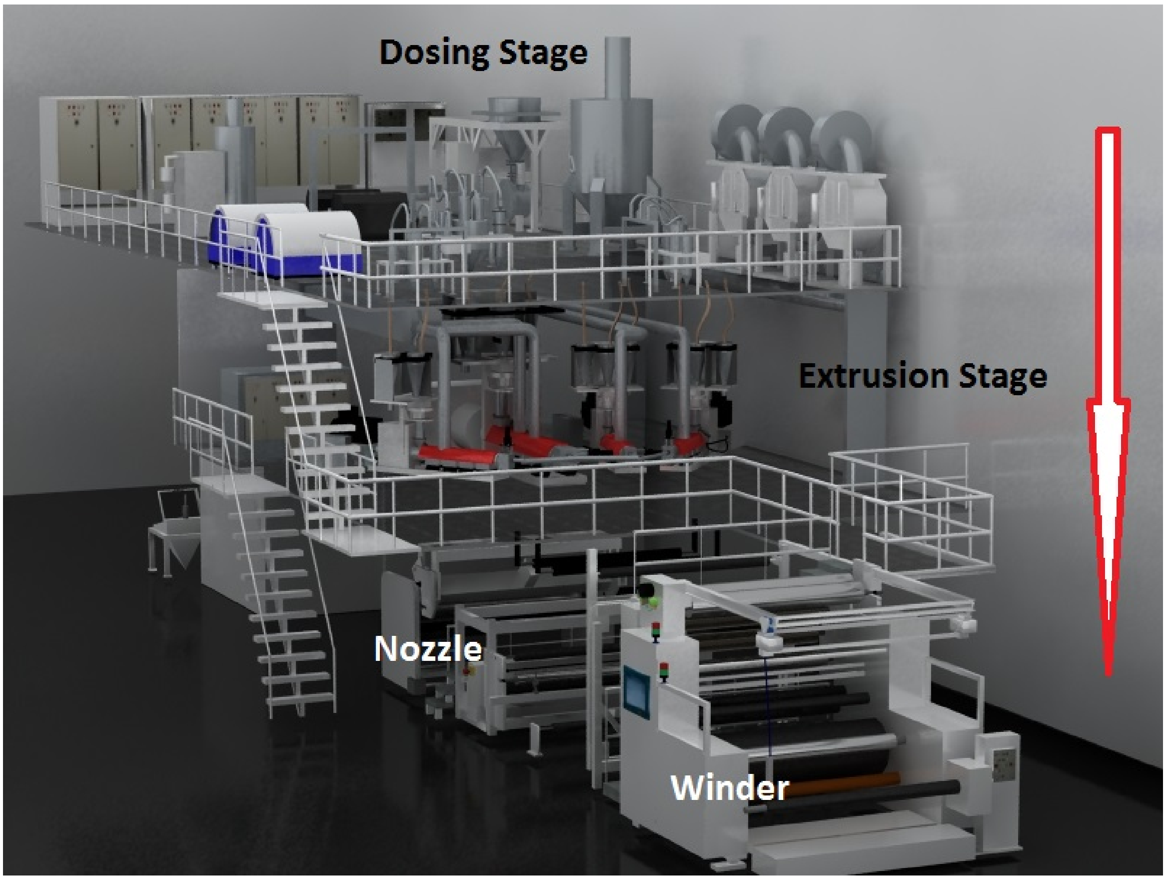


Figure 16 Cast extrusion process consisting of heated dosing stage, various extruders, nozzle, and a winder with slitting [K+10]

The main processing part is the extruder (containing 1, 2, or more screw conveyors), which is charged with granulates heated up at 170–220°C, generating a melt that is pushed out by the screw conveyor; an example is displayed in Fig. 17. Several attributes have a high impact on the production quality, as e.g., the pressure, which needs to be kept low [Gne14].

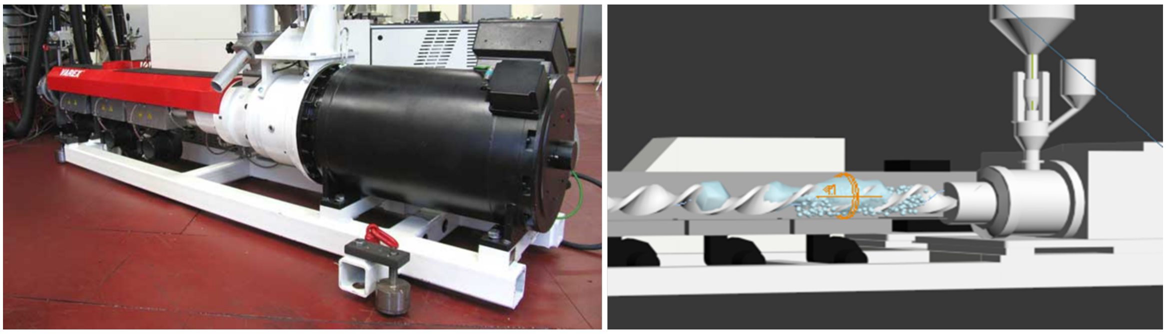


Figure 17 Cast extrusion screw conveyor animation filled with PE material, at 170-220°C

Thickness measurement and optical control systems for in-line quality control are usually additionally integrated in such lines at the winder. About five PLCs (e.g., S5, S7) regulate the machine behavior, such as temperature and speed, through up to two operators [S^+14].

Blownfilm-Extrusion

The blown film extrusion usually has a lower dosing stage feeding the extruders with, e.g., PE. The molten granulate (190°C) is injected into a ring that forms the blow (warm temperature inside of about 60°C, cooled from outside with 20°C), which moves up to a traversal. There, the blow is pressed to a flat film by traversal rotation and then guided to the winder below for rolling up, as shown in Fig. 18. The rolls are slitted in a second additional step, if necessary.

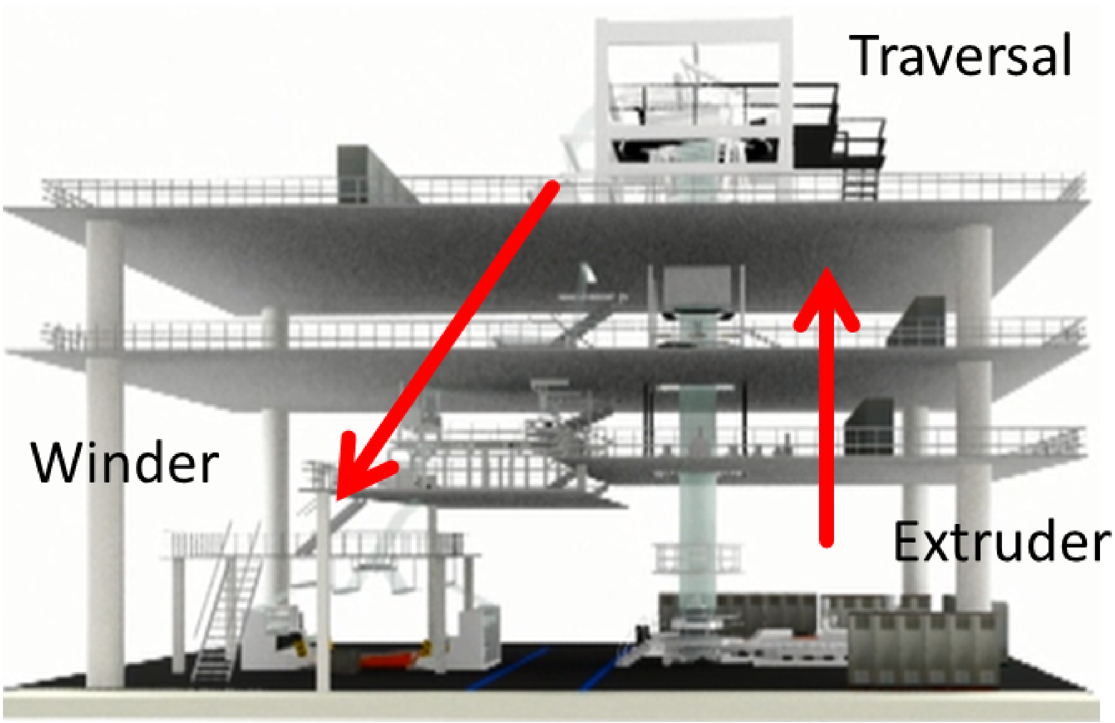


Figure 18 Blownfilm extrusion process consisting of extrusion stage (e.g., 3 extr.), the traversal, and a winder without slitting [K^+10]

The blown film extrusion process has low material waste and stable process conditions, in comparison to cast extrusion. A regular machine generates about 600–1000 kg output per hour (10–20 to per day), with an energy consumption of 250 kW/h (efficiency 0.6 kW/h/kg). About five PLCs (e.g., S5, S7) regulate the machine behavior, such as temperature and speed, through up to two operators. [S+14]

Calendering

The calendering[13] process starts with a heated dosing stage (up to 120°C) for elimination of humidity and residual monomers, followed by mixing machines for cooling (60-40°C) of, e.g., PVC powder combined with additives (from the silo or big bag). The mixed batch moves into a kneader (180°C) where it is melted up, and after feeding into the calender (e.g., L-Type) (200°C), it is transformed into a thin film (200 m/min), rolled up and slitted on a winder[14], as presented in Fig. 19.

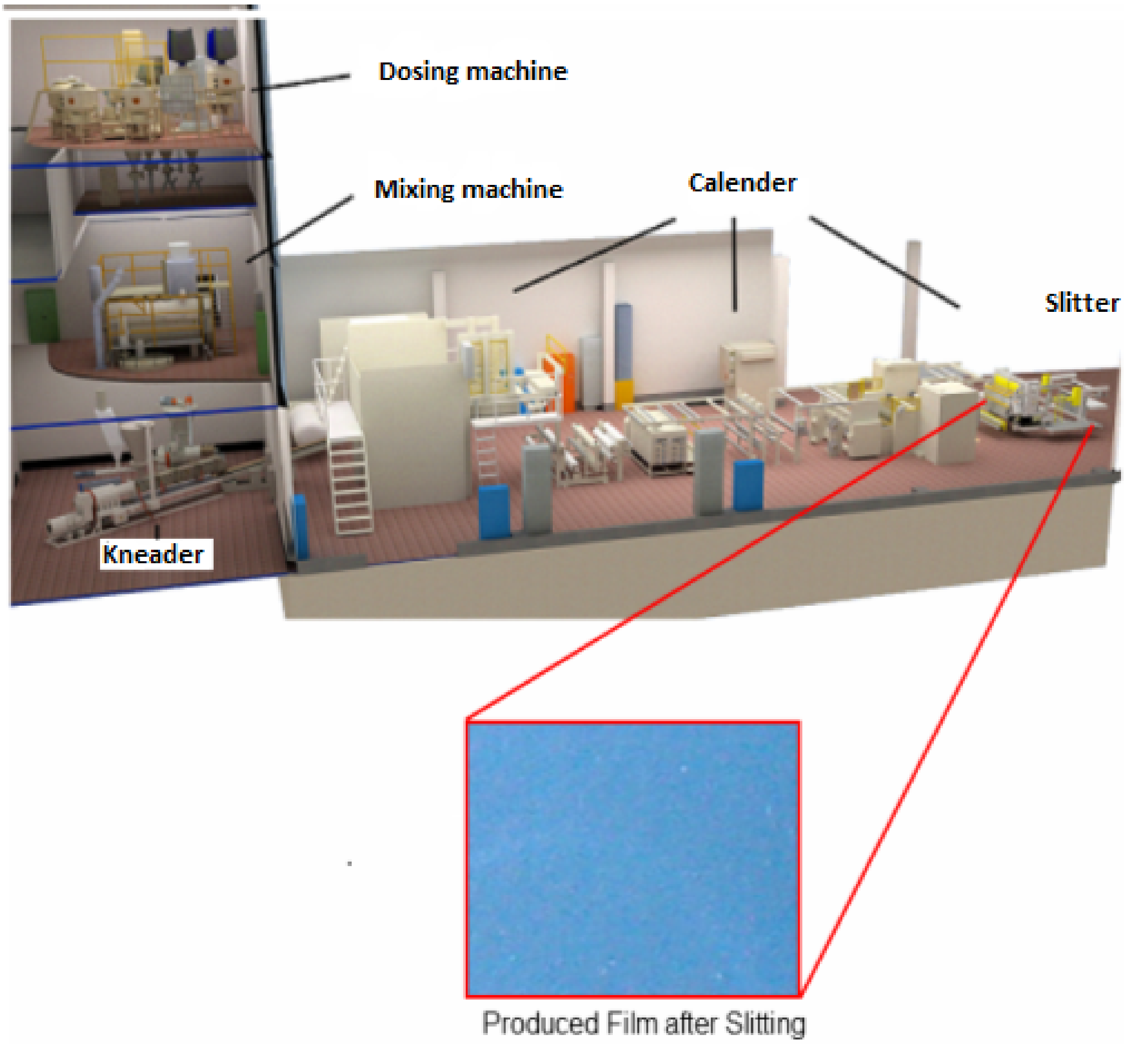


Figure 19 Calendering process consisting of heated, and cooled mixing stages, the kneader, the calendering line, and a winder with slitting [K+10]

Such process lines (Fig. 19) include about five PLCs (e.g., S5, S7) operated from a master active control station and, in part, manually by operators at the line. Quality measurement is managed off-line in laboratories or in-line by sensor control systems (e.g., thickness measurement or optical control systems[15] [[K+92] [S+14] [OCS15]]. The

[13] Handbuch Urformen
[14] Winders exist with or without slitting
[15] Reflected, or transmission light systems by e.g., OCS GmbH (2015)

calendering advantage compared to extrusion lines is, higher productivity of up to 4 to/ hour, and faster switching from different thicknesses between 50 to 1200 μm. In contrast the disadvantages are nearly twice the number of capital investment, and slower set-up time for colour changes.

3.1.3 Electronic data processing

Manufacturing plants possess different data sources, such as manufacturing execution systems (e.g., Coago), enterprise resource planning (e.g., SAP/CRM, SCM), machine databases, quality databases (e.g., Busitech), sensor databases (e.g., OCS GmbH) and others, which are connected within a TCP/IP network. Separated into in-line and off-line data acquisition systems, they are accessible at different parts of the plant through interfaces. Figure 20 gives a holistic overview of the data sources within manufacturing plants, here taken from the polymer film industry as an example, regarding Fig. 6.

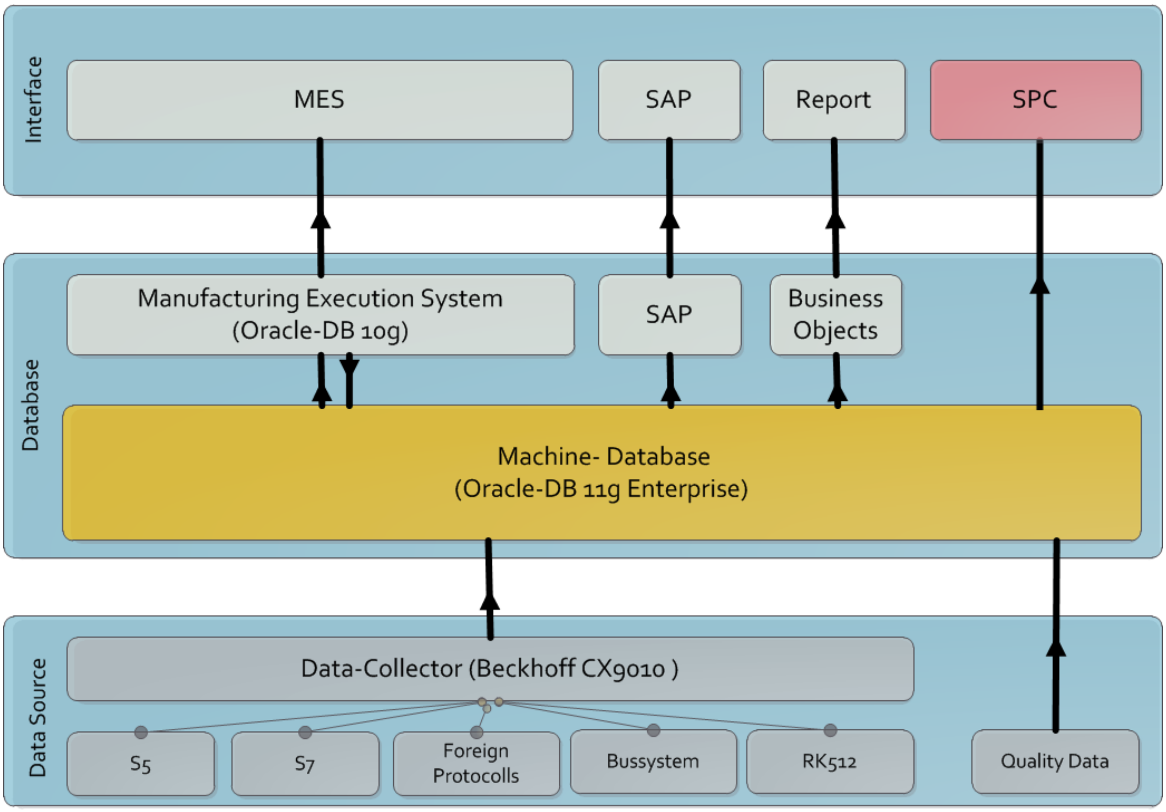


Figure 20 Exemplary database framework within polymer manufacturing plant

Beginning with the data sources, different foreign protocols and PLCs offer datasets for data collection and conversion into databases. Sensor location points at extrusion lines (e.g., Speed, Thickness, MaterialType) and manual operator input fill the above-mentioned data storages with information. Monitoring systems allow in-line process surveillance of specific datasets, e.g., speed, temperature at the machine terminals, adjusted by experienced machine drivers. Tools as shown in Fig. 21 allow statistical post-analysis.

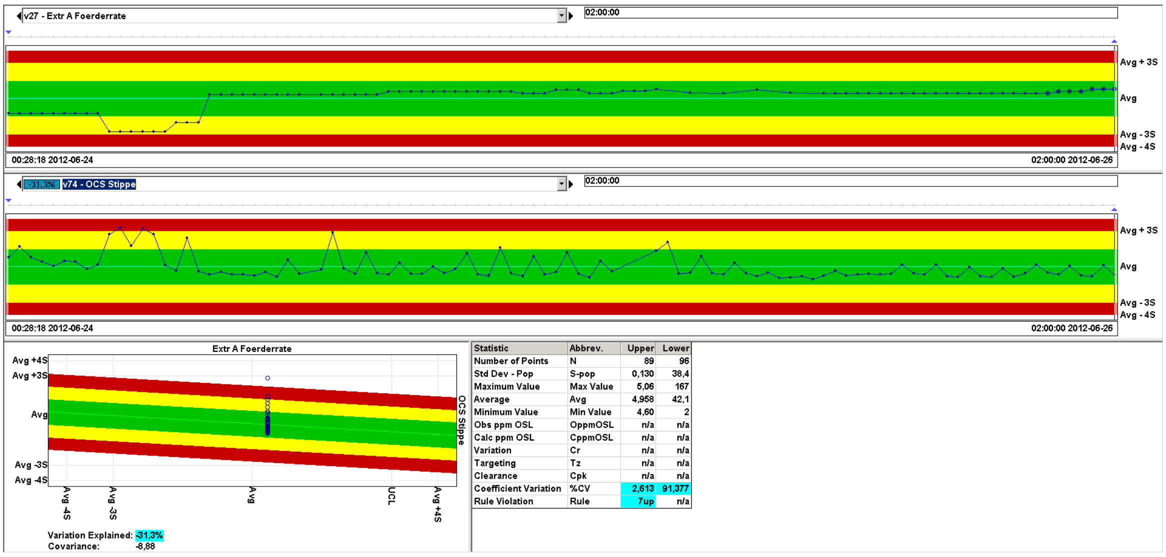


Figure 21 The *Quality Window* software interface for statistical process control [Bus15]

A statistical process control (e.g., *Quality Window* by Busitech; Fig. 21) is integrated in the surveillance of process deviations by quality departments, checking for customer and raw material specifications. Besides, Minitab[16], Excel or specific reporting tools (e.g., *Business Objects*[17]) are used to prevent compensation claims due to quality problems. Such tools can be integrated for in-line monitoring of production, but this is usually not done in smaller plants. For this purpose, sensor applications are installed for specific attribute monitoring, such as optical control systems (defects: gels, holes, contaminations) and color (L*a*b) or thickness measurements (μm).

3.2 Process Data

Process data in the manufacturing industries contain data from sensor locations at machines, readable via PLCs, and quality data from separate monitoring systems (e.g., optical control systems) and off-line laboratory systems. In contrast, data on process specifications for recipe settings, parameter limits, or customer requirements come from different sources and are stored in different formats prepared with QuickCog feature selection methods (e.g., sequential backward selection (sbs), sequential forward selection (sfs), qs separability, RNN) for, e.g., cause seeking (relevance), to be monitored downstream by analytical systems.
The following Subsections describe the different sensor locations (e.g., dosing status in binary format) at specific polymer machines, their settings, ranges, drifts, and failure rates, compare these with quality in-line and off-line laboratory measurement applications, and conclude with process specifications depending on different products and their processing environment.

3.2.1 Sensory Locations (Settings, Range, Drift, Failure rate)

Polymer manufacturing machines, equipped with hundreds of sensors from the dosing stage to the winder stage, collect heterogeneous sensor data (e.g., on temperature,

[16] Minitab Inc.
[17] SAP

pressure, speed, or energy consumption) from the machine controllers (e.g., S7) via TCP-IP, as shown in Fig. 22. Within each production minute, sensor data is stored on database servers in specific format types (e.g., numerical or char values) validated by technical experts for further downstream process analysis.

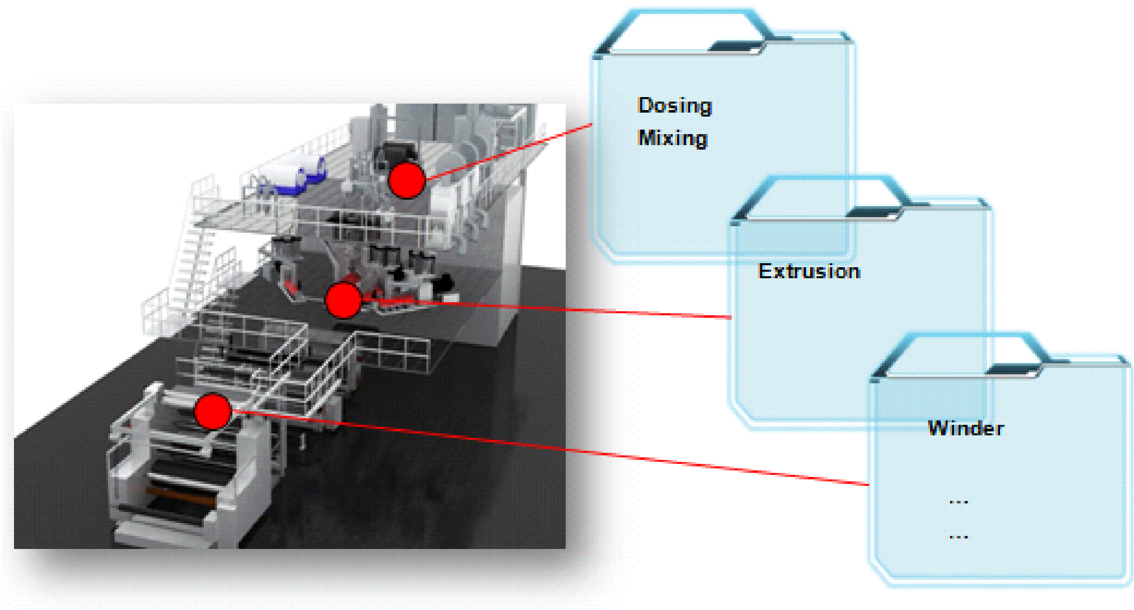


Figure 22 Sensory points at typical cast-extrusion line from dosing to winder stage

Calendering and extrusion machines, depending on their PLC layout, allow the acquisition of up to 2000 process measurement values by multi-sensor modules at fixed processing points. The sensory measurement values are collected by multiple sensors, e.g., from the *Temperature* zones, the *Melt pressure* parts, and the *Power main drive* energy, groupwise (2–10 sensors) or singly located over distances of 1–3 m, as displayed in Fig. 23.

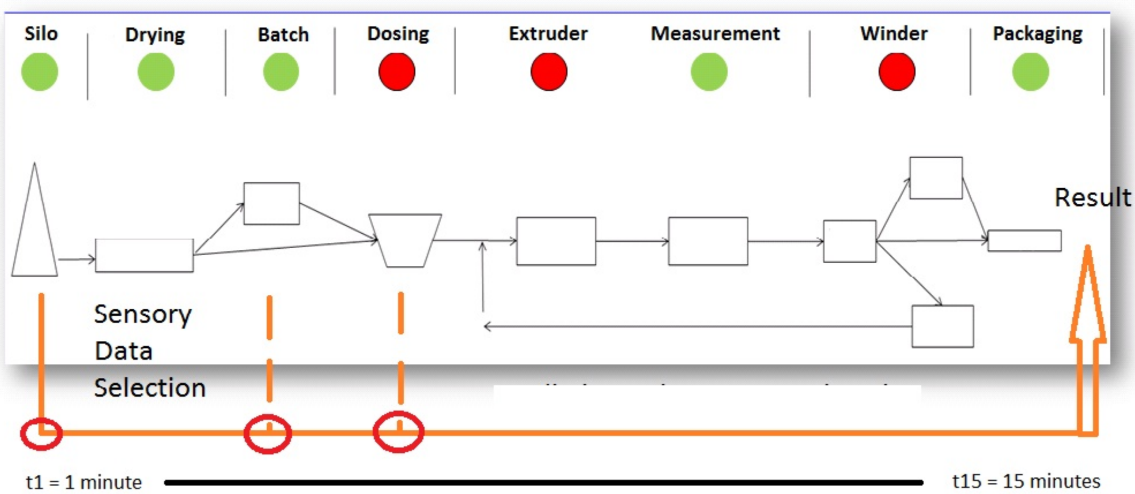


Figure 23 Sensory locations distributions at typical cast-extrusion line from dosing to winder stage

In Figure 23, the sensor locations are shown for a typical extrusion line. Within 15 min, each kilogram of polymer passes about 160 sensor locations, which are unequally distributed (see Fig. 24).

Elapsed time range	Number of sensory locations
t1 = 1 minute	1:5
t2 = 2 minutes	1:10
t3 = 3 minutes	1:30
…	…
t15 = 15 minutes	1:160

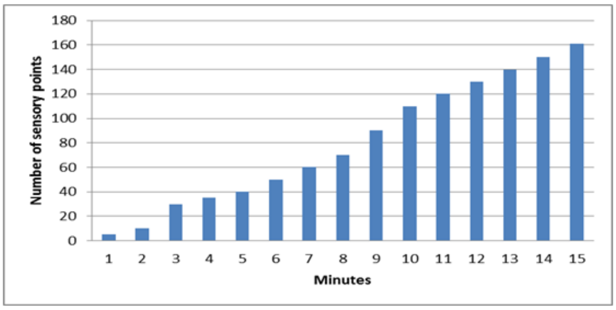


Figure 24 Selected sensory locations quantity depending on elapsed time (typical extrusion line)

Today's polymer extrusion machines are able to acquire data from up to 2000 sensor points in the process, due to higher automation standards and improved PLC types. The ranges of some typical sensors from the extrusion process are shown in Table 2. Location induced failures of sensors, e.g., pressure or temperature sensors, within 2 years are caused by technical issues (Hodge (2004) [HA04]) or environmental influences, leading to faulty outputs coming from transmission or dataset generation problems. The transferred data format contains integer, real, float, Boolean, or categorical values related to the sensor settings for data export.

Table 2 Ranges of sample sensors from extrusion line process according to Gneuss (2015) [Gne14]

Sample Sensors	Type	Ranges	Accuracy	Sample Rate	Resolution
Temperature	PT100 FeCuNi DTAI Series	max. 400°C, max. 2000 bar	0.5 %	120 ms	-
Pressure	DAI Series	max. 400°C, max. 2000 bar	0.5 %	-	16 bit

3.2.2 Quality Data

Present-day polymer quality departments need to check the second or third roll in the product series of an order, with regard to process instructions and customer requirements. Typical attributes measured are thickness (deviations), strength, gels (>2 mm), peak load (5%, 10%), color (L*a*b), and additional specifications. The data is typically stored in Excel and Minitab or Quality Window[18] and other data systems for off-line local laboratory analysis. On the other hand, recognition systems for on-line quality monitoring integrated into machines are *State-of-the-art* for optical control (Fig. 25), thickness, color, and gloss measurements.

[18] Busitech

Figure 25 Optical transmission light control system (OCS GmbH) for defect detection in polymer films

The transmission light optical control system by the OCS GmbH from Fig. 25 monitores defects, as gels and holes in real-time at the Mondi Gronau GmbH. About 14 systems are integrated at the plant directly connected to the firm-internal ethernet network. Normally, the acquired data of all types are stored locally on standardized industrial computers, directly at the production line (see Figure 3). Quality data is typically

Table 3 Laboratory data for specific orders regarding time, output, thickness, and additional information from Quality Window

Order	Time	Output [kg]	Thickness [μm]	...
20003344	15.01.2014 12:00	200	54.1	...
20003344	15.01.2014 12:05	198	54.5	...
20003345	15.01.2014 12:10	235	58.2	...
20003345	15.01.2014 12:15	230	58.2	...
20003345	15.01.2014 12:20	234	58.9	...

associated with order numbers, time range, customer ID, machine number, and other specifications, combined in tables related to the monitored attributes, e.g., thickness, output, and more.

3.2.3 Process Specifications

The process specifications[19] to fabricate specific products are called recipes for machine settings, including about 30 adjustable parameters such as temperature zones (e.g., 1–10) for extrusion lines, speed (m/min) regarding the necessary output (kg, m^2), status settings for the dosing ratio (1–100%), winder cutting points if slitting is required, and activation of measurement systems for defect detection, thickness and color monitoring,

[19] Manufacturing Instructions

as in part presented in Table 4. Besides this, there are customer requirements including

Table 4 Extracted partial settings from process specification for cast-extrusion line elastomere product

Parameter	Unit	Min	Target	Max
Thickness	μm	71.0	79.0	80.0
Width	mm	195	205.0	220
Defects	#	0	100	120
Knife Position	m	2.4	2.5	2.6
Aspiration	%	60.0	100.0	100.0

targets for quality data such as thickness ranges (e.g., 40 μm $\pm$ 10%), the number of holes (e.g., max. count of 5 of size 2 mm++ on 8.000 m roll length), or the material ratio within film layers (e.g., inner to outer layers 2.5%/95%/2.5%).

3.3 Challenge

Modern polymer manufacturers are confronted with the consequences of globalization (international competition), high cost of raw materials or energy, lower margins, high technological complexity resulting from processes involving various stages, additional aggregates, monthly changing extremely sensitive products, smaller lots due to just-in-time production, multitudes of different machines, precise (μm) modifications needed, lack of time for and knowledge on real-time analysis, decentralized, insufficiently linked heterogeneous data storage devices, and a lack of monitoring systems satisfying the needs of machine operators, quality employees, and engineers. Only 5% of the possible manufacturing data is used for error cause analysis due to missing data source linkage and insufficient knowledge on data mining methods applied to huge datasets. Therefore, the resulting conclusions from production, quality, and engineering surveillance are often biased. Increasing losses of money in the range of billions of dollars/euros are caused by the previously mentioned problems leading to material waste, customer claims for damages, safety issues, and high energy consumption.
This Chapter aims to select three of the above-mentioned topics with strong influence on material waste, and energy losses and to apply them to a polymer film industry case study: material waste, process monitoring, and energy efficiency. Waste reduction with focus on a set of extrusion raw data attributes (focused sensor types; typical ranges shown in Table 2), whereas process monitoring shall build up on previously found results with extracted extrusion raw data attributes, allowing an excursion on energy efficiency with multiple attributes of energy consumption data (displayed in Table 2).

3.3.1 Waste Reduction and Yield Optimization

According to Plastics Europe, and Consultic[20] (2013) the reduction of landfill disposal (Fig. 26) showed a positive trend in 2012 of about 5.5%, remaining at the same level of 25.2 million tons of mainly packaging since 2011 (77% from Germany, the UK, France, Italy, Spain, Poland, and The Netherlands (EU-27+N/CH)) [Pla13].

[20] Marketing & Industrieberatung GmbH

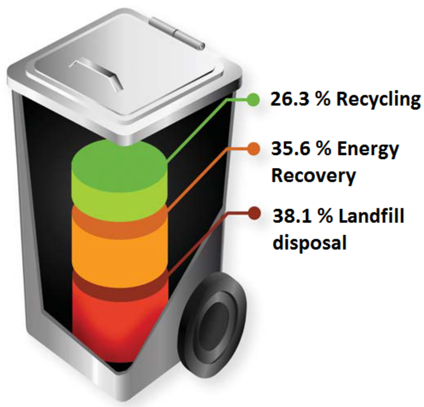


Figure 26 Disposal, recycling and energy recovery in 2012 according to Consultic [Pla13]

Success in improving waste reduction strongly depends on knowledge of the actual process behavior, which is not always given in complex extrusion production. In multi-stage manufacturing processes it is highly recommended to improve earlier intervention in machine settings to maintain quality assurance. The final product cost for material waste after a three-stage process can be reduced by shut down within the first stages, as shown exemplary in Fig. 27. The material and staff cost generated per minute over three stages are displayed. The waste material for a three stage compounder, extrusion, and lamination process can be decreased by predictive intervention in the first compounder, or the second extruder stage. The earlier deviating conditions are recognized, the faster counteractions can be initiated and cost reduced.

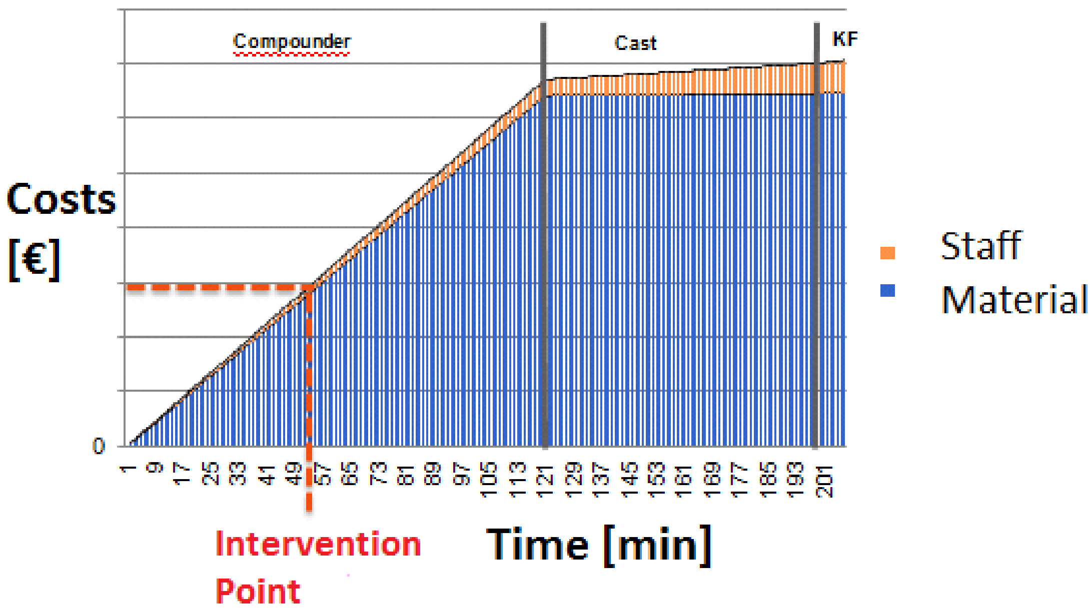


Figure 27 Exemplary intervention point for predictive waste reduction within three-stage polymer process

Therefore, Chapter 4 will examine different process condition datasets with unknown situations leading to material waste (called landfill disposal in Fig. 26). A research approach shall be done with modified classification methods on different datasets with multiple attributes from rigid film raw data.

3.3.2 Process Control and Visualization

Process condition and sensor failures are not randomly upcoming events but, in most cases, environmentally influenced parts of a chain of events heading towards an abnormal event. Such states show pre-events in earlier time ranges, hinting at the main problem. Regarding the prediction of process conditions and sensor data deviations described in the previous Subsection, improvement on faster recognition could be achieved by trajectory behavior analysis as presented in Fig. 28.

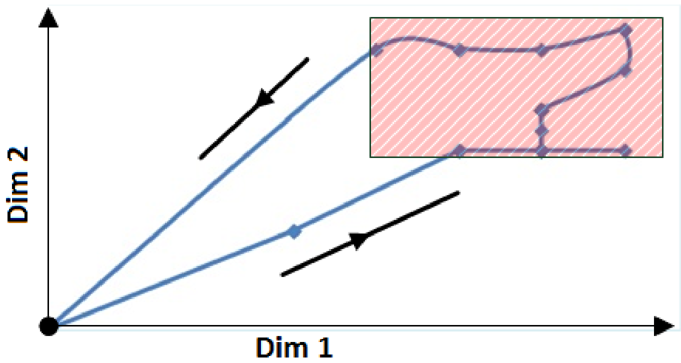


Figure 28 Exemplary trajectory visualization of polymer process data (Dim 1 vs. Dim 2, critical part shaded)

A typical visualized trajectory from Fig.28 could run into a critical range, giving information on a condition change resulting in material waste due to process problems. The visualization of the polymer process data and its separation into time shifts for trajectory analysis will be shown for real-time extrusion data based on selected attributes from extrusion pressure raw data. The research approach in the following Chapters will aim at the recognition of events several minutes before their actual occurrence, to allow for adaptations of the process settings.

3.3.3 Prediction and Reduction of Energy Consumption

In Germany, production plants with unavoidable high energy consumption are able to request an energy net release fee[21] at the end of the business year, depending on calculations of the relation between the yearly energy consumption (kilowatt hour [abbr.: kWh]) and the highest peak load (kilowatt [abbr.: kW]). The energy consumption represents the sum, whereas the peak load the maximum performance of all 15-minute-intervalls during the year.

[21] Bundesnetzagentur, Institution for energy net stability, [Bun13]

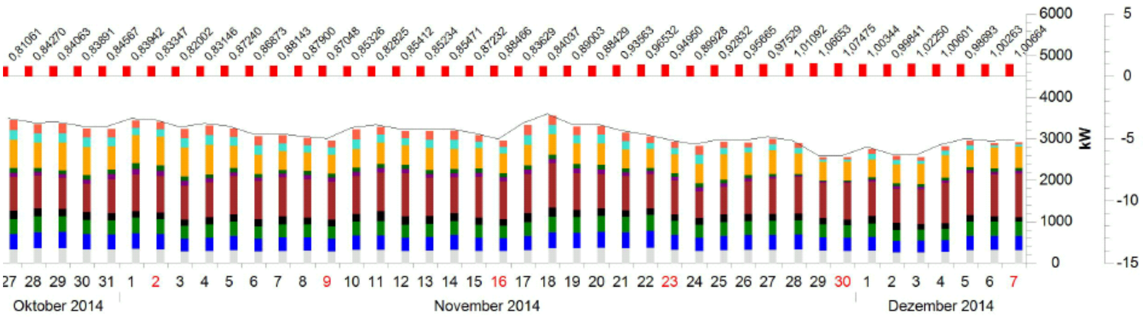


Figure 29 Extraction of peak loads from transformers at the Mondi Gronau GmbH in 2014

Figure 29 shows the peak load behaviour for all transformers at the Mondi Gronau GmbH at the end of 2014. The net stability ratio is defined in the following calculation:

$$\frac{\sum Annually\ Energy\ Consumption}{Maximum\ Peak\ Load} > 7000\ h \tag{2}$$

By maintaining 7000 hours, and more, the energy consumption is kept in relation to the peak load at a stable level, and the annual energy cost could be lowered by up to 11 %. This way, steady energy consumers are supported by the Bundesnetzagentur.

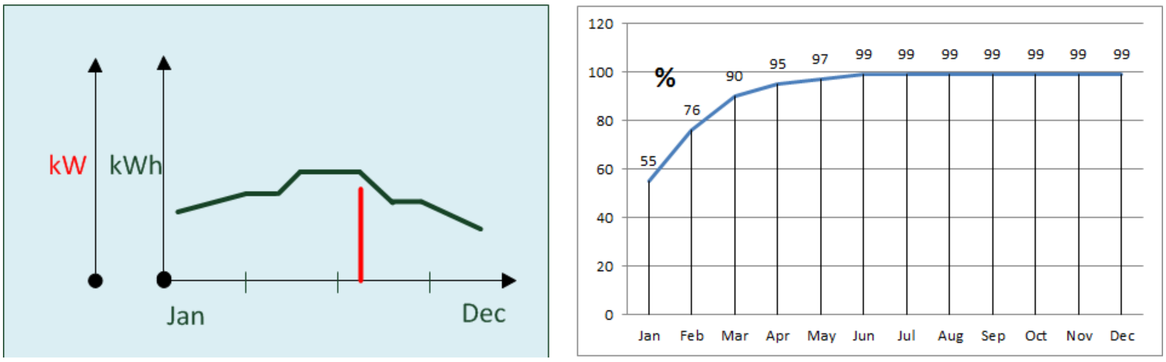


Figure 30 The energy consumption, maximum peak load, and prediction accuracy within typical business year for specific polymer plant

The main problem of polymer companies is that they cannot predict their final amount of energy consumption and their highest peak load, due to market demand changes (regarding short forecasts of 1–2 weeks) during the whole year, as shown in Fig. 30. Thus, they are advised to use energy forecasts necessary for controller calculations several months before the end of the business year.
Novel approaches are able to develop models using training data from the past years to predict the energy consumption behavior depending on seasons and environmental influences. Chapters 5, and 7 will present a case-based research for predicting the energy behavior in a high-consumption polymer production plant, based on multiple attributes of monthly consumption data from several years.

3.4 Case Study Mondi Gronau GmbH

The Mondi Gronau GmbH is a leading international supplier of various films and film-based products. The product range includes high-quality printed packaging solutions and films, technical films, hygiene components, laminating films and label films as well as decorative films [Mon15]. As leading polymer manufacturer for thin multilayer films for the hygienic, automotive, and other industries worldwide, about 65 different producing and converting machines from extrusion and lamination to offset printing are located there.
In the following, the lowest structural part of the complete Mondi corporate group is described by the machine object, Fig. 31.

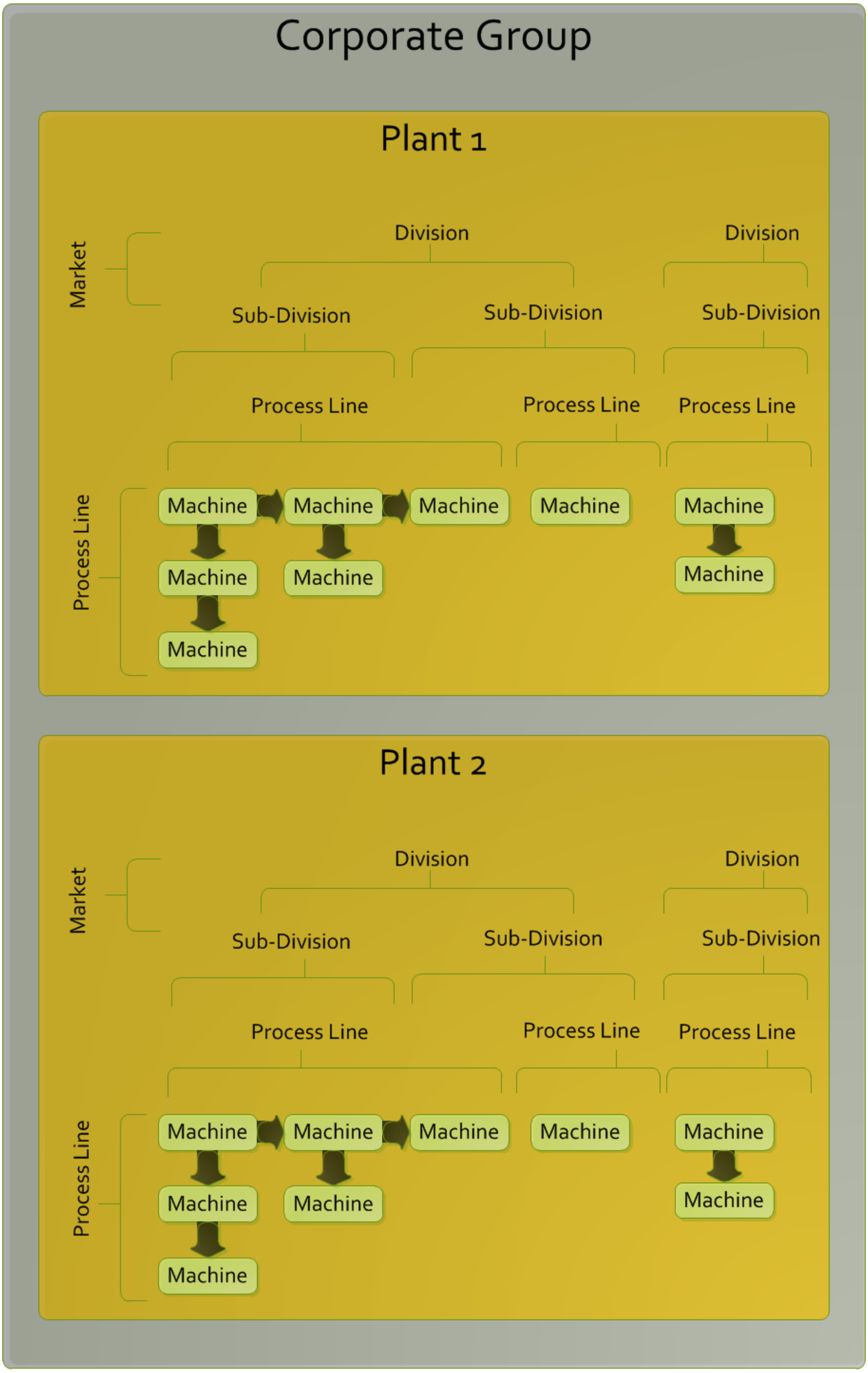


Figure 31 Typical polymer film corporate group overview with objects, as machine, process line, stage, division, plant, and market

The process lines above consist of 1 – 3 consecutive machines, integrated into stages beside including connected process lines. The divisions, or business units, are supervising several process lines, and stages, connected to the market via distribution. The previous parts represent the plant, displayed in Fig. 32, and connected to other plants the corporate group.

Figure 32 The Mondi Gronau GmbH virtual overview

For general validity, the applied research approach will be employed at the main polymer plant in Gronau (Westf.), regarding waste, monitoring, energy, and recognition issues.

4 Anomaly and Novelty Detection

Raw-material property changes, deviations in process condition behavior (e.g., due to environmental influences), and sensor problems (e.g., transmission failure) lead to increased processing cost and lower yields. Novelty detection and anomaly filtering are of high importance in the manufacturing industries when observations are not properly sampled and the limits of the operators' experience are exceeded. A particular, novel approach in the field of classification is *One-Class Classification* (OCC) [MH96], allowing predictions on sparsely sampled process datasets for unknown objects. This can be exploited for process control to avoid yield losses.
The trained OCC offers timely recognition of process abnormalities from multiple sensor locations and preprocessed additional features with a low amount of information, whereas the accuracy of multi-class approaches, as explained in Chapter 2, is at its limit [KM10]. This Chapter investigates *State-of-the-art* classification methods, regarding their adjustment to OCC issues. The modification of classifiers to *One-Class* approaches offers advanced recognition abilities for sensor, and condition deviations.
In the following, Section 4.1 explains the fields of application for OCC. Section 4.2 describes the OCC methods, focusing on novelty recognition in the polymer industry. An overview of the conducted research experiment and an evaluation of the methods are presented in Section 4.3. The results are discussed in Subsection 4.3.3, whereas Section 4.4 provides a summary and outlines future work.

4.1 Overview of Applications

Real situations where only one group of data is presented occur in different parts of life: novel system states in nuclear plants [Tax01] leading to dangerous incidents, network intrusion detection for companies to protect data from misuse, biomedical recognition of cancer and diseases in their early stages, signal processing, computer vision systems, bank frauds, and many more [[T+95] [CM95] [B+11] [RG07] [D+10] [Bar10] [Bor09]].
Support vector machines, k-nearest-neighbor classifier, and neural networks approaches in, e.g., image data processing are based on predefined adequate knowledge of all possible or nearly all possible conditions needed by *State-of-the-art* multi-class classifiers. Spam filtering or intrusion detection systems face unknown events which cannot be handled by multi-class assignments. Therefore, novelty one-class approaches for anomaly detection adjusted to different processes deal with missing (unknown operating conditions) and faulty (dummy values from PLC) data to increase the detection accuracy, as described by Tax et al. (2004) [[S+05c] [TD04]].
The approaches in these fields are distinguished into parametric and non-parametric types.

4.2 Clustering and Classification

Clustering (unsupervised learning) aims at grouping objects that are similar to each other into clusters, by clustering algorithms. The distance to a cluster group decides about the assignment of an object. Different models for clustering can be used, e.g., centroid or distribution models for a particular problem. Their accuracy or cluster validation measure is a criterion for similarity when comparing chosen models [Bai94].
Classification methods (supervised learning) identify the correct assignment to a particular trained category for new observations. The performance of classification systems

depending on the given datasets is evaluated by receiver operating characteristic (ROC) curves, which define the false-positive rate of observations [Mit97].

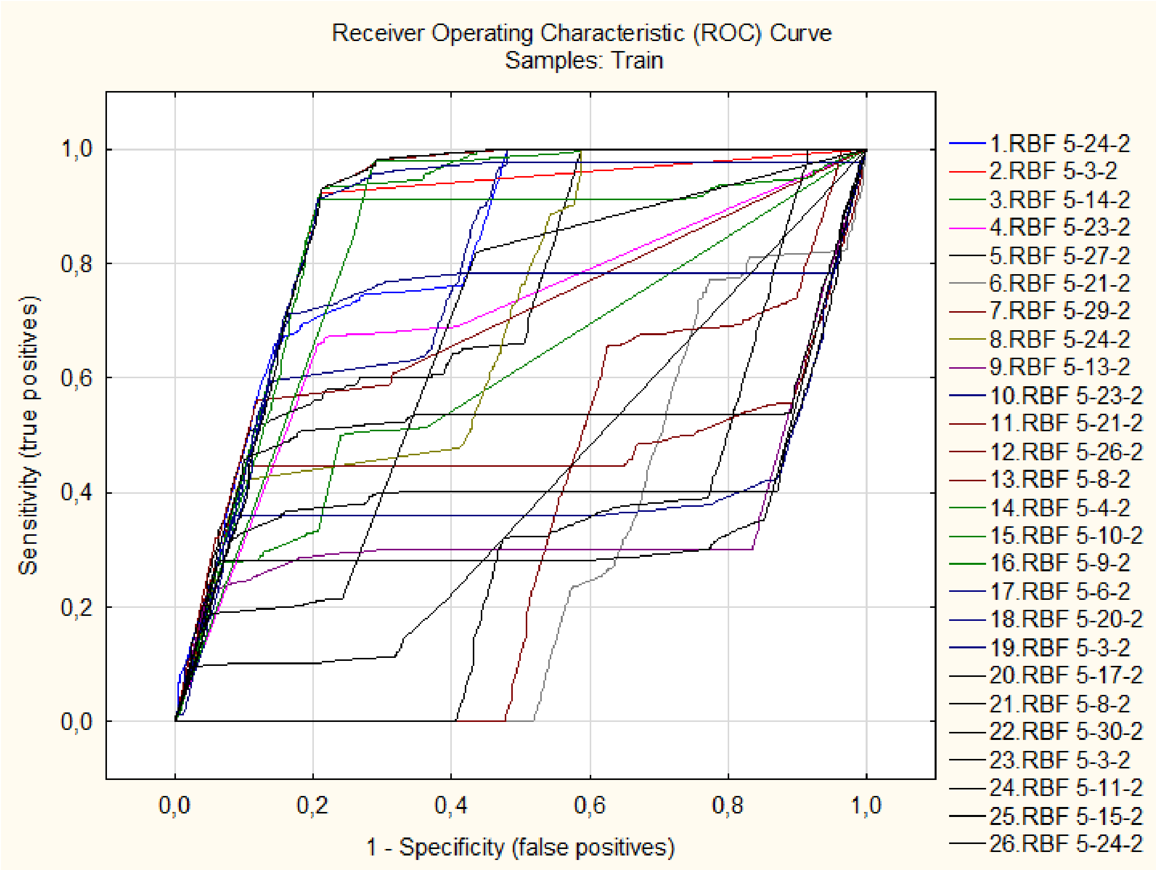


Figure 33 Receiver operating characteristic curve for 26 rbf neural networks based on real production data from polymer industry

The ROC curve for 26 radial basis function (RBF) neural networks is presented in Fig. 33. The following Subsections describe in more detail the chosen exemplary classification methods with regard to their actuality in manufacturing applications. Methods like neural networks, support vector machines, decision trees, or k-nearest neighbor classifier (e.g., optical control systems) are *State-of-the-art* in the field of classification approaches [Gna97]. The present-day optical control systems for visual defect classification use decision trees, support vector machines and k-nearest-neighbor methods, whereas neural network methods are preferred in predictive analytics.
The following methods from supervised learning were chosen accordingly to the listing by Khan (2010), who summarized classification types by authors (e.g., Manevitz and Yousef (2001), Tax, Duin, Ridder (1998), Schölkopf (1999)) regarding their behaviour in sparsely sampled classification approaches, Section 2.4. In such cases the methods were modified for delimiting the contour of the observations by boundaries [[Tax01] [MH96] [KM10] [S+99] [MY01] [R+98]].

4.2.1 Neural Networks

Artificial neural networks mimicking central nervous functionalities with neurons (nodes) and weighted node connections within a network represent the non-linear functions of their inputs. Used for supervised, unsupervised and reinforcement learning, artificial neural networks infer functions from observations for complex tasks [[Bis95] [Hay98] [CU93]]. In Pseudo-Code 1 a typical procedure for the backpropagation neural network is described.

begin

1. Create architecture of neural network (*e.g.*, $4-7-2$)
2. Randomly generate initial weights
3. If the error is to high perform for each training pattern
   3.1 Calculation of every neuron (input-hidden-output layer)
   3.2 Calculate error at outputs (e.g., cross-entropy error)
   3.3 Error signals for weight adjustment
4. Repeat evaluation by test set for performance ratio

end

Pseudo-Code 1: Procedure of a basic neural network classifier with weight adjustment after cross-entropy error calculation

The parameter settings are adjustable by the selection of the activation function, the learning rate, the regularized coefficient, the number of units, and the number of hidden layers. Furthermore the *Auto-Encoder Neural Network* is an unsupervised learning method extension of the neural network classifier.

Automated Network Search - Multilayer Perceptron
The Multilayer Perceptron (MLP) is a feedforward artificial neural network with several layers using a nonlinear activation function for each neuron [Hay98].

4.2.2 Support Vector Machines

Support vector machines (SVM) are supervised learning algorithms for training models of two categories for real-world problems with linear or nonlinear classification (well-known for pattern recognition), by constructing one or more hyperplanes in a high-dimensional feature space. A sample SVM for pigment recognition in polymer films by optical sensors for counterfeit detection is shown in Fig. 34 [[CST99] [CST00] [Bur98]].

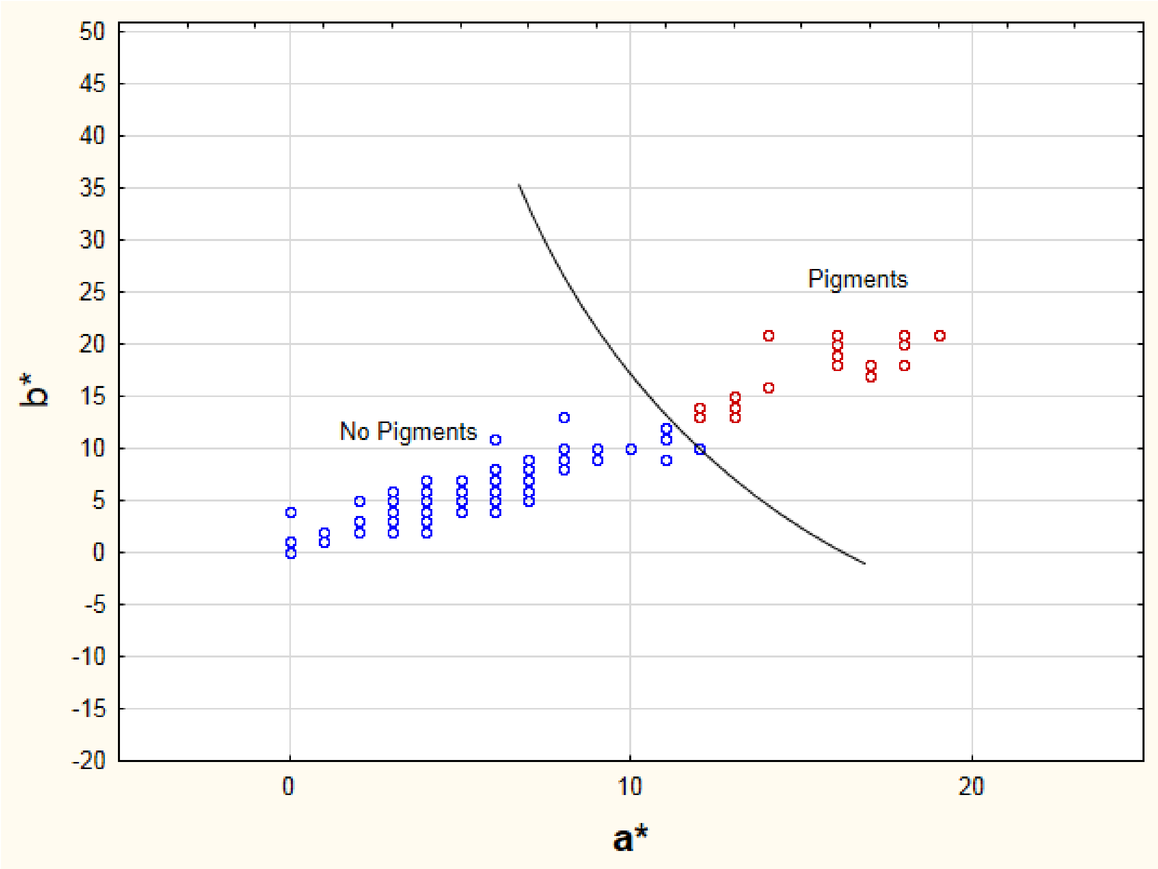


Figure 34 SVM trained on real optical sensory data (L*a*b Values) for pigment recognition in polymer films

The margin defines the performance of observation classification with a kernel function. General application fields for SVM algorithms are the medical sciences and image and text classification issues [CV95]. Novel approaches in fault detection combine genetic algorithms and support vector machines [JN02]. An exemplary algorithm sequence is displayed in Pseudo-Code 2.

begin

1. Define $\mathbf{X}$ as input and $\mathbf{Y}$ as Label
2. Categorization in two or more classes $\mathbf{Y} = (1, -1)$ or $\mathbf{Y} = (1, ..., n)$
3. Training dataset $\mathbf{S} = ((\mathbf{x}_1, \mathbf{y}_1), ..., (\mathbf{x}_t, \mathbf{y}_t)) \subseteq (\mathbf{X} \times \mathbf{Y})^T$
4. Sample distance to hyperplane $d = y_i(\langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b)$
5. Compare d_i with d_{i+1}
6. Selection of maximum margin hyperplane d
7. Repeat until best result found

end

Pseudo-Code 2: Procedure of a support vector machine classifier for hyperplane margin

Y is a column vector with as many rows as the complete matrix X, and the sample matrix S, inside both each row corresponding to an observation. t describes the number of observations as a counter. $< w, x >$ is inner (dot) product (scalar) of w and x, with w as the normal vector, x as the observation vector, b as the bias.

4.2.3 Decision Trees

Decision trees from the field of supervised learning for decision analysis use rules to divide datasets for classification or regression purposes. Decision trees have been extended to more complex algorithms for solving high-dimensional problems. They are often used for solving operations research problems. Their flow chart handling makes them useful for decision rules, as shown in Fig. 35.

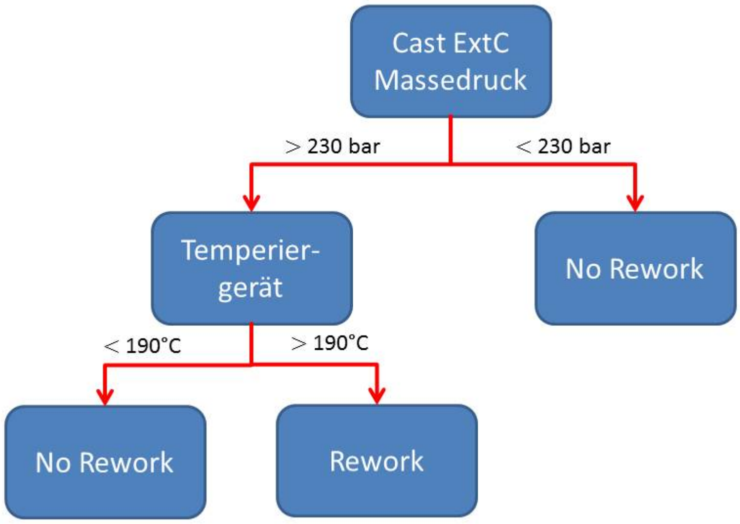


Figure 35 Tree graph sample for rework state on real polymer production process datasets made with STATISTICA

In Figure 35 a decision tree for rework conditions based on real polymer production data is displayed, consisting of 6 non-terminal and 7 terminal nodes. Each node represents a decision rule depending on the dimensional size of the dataset. Additionally newer types like "boosted trees" are becoming more popular.
As shown in Pseudo-Code 3 the attributes of labelled datasets are chosen that split the set of objects into smaller subsets, which is called recursive partitioning, combining the quality of splitting with a performance measure (information gain). [Qui86]

begin

1. Labelled Data = $(\mathbf{x}_1, c_1), \ldots, (\mathbf{x}_n, c_n)$
2. Instance attributes = $(x_{1,i}, \ldots x_{n,i})$
3. Choosing attribute x_{ki} with highest information gain to split group

   Normalized information gain ratio

   $$IG(\mathbf{S}, x_{ki}) = H(\mathbf{S}) - \sum_{t \in T} p(t)H(t)$$

   $\mathbf{S}$ = set of training samples

4. Create decision node at attribute x_{ki} splitting into subsets
5. Repeat procedure for subsets to generate children of previous node

end

Pseudo-Code 3: Procedure of a decision tree classifier with normalized information gain ratio

4.2.4 *k*-Nearest-Neighbor Classifier

The k-nearest-neighbor classifier (kNN) methods from the field of supervised machine learning train models of different classes to be used on new datasets for specified classification. kNN is a non-parametric simple classifier categorizing a new observation by majority vote into a class. The k input variable by kNN defines the included number of neighbors (more specifically, the class labels) and their distance weight for multi-dimensional classification.

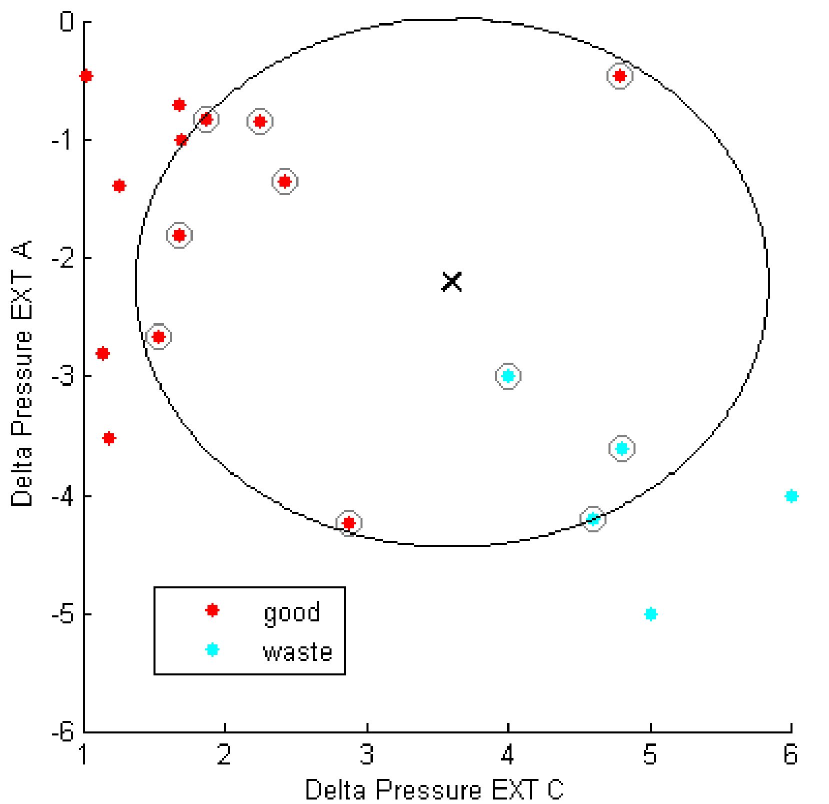


Figure 36 k-nearest-neighbor classifier trained on real manufacturing pressure data (waste condition, good condition)

If 10-nearest-neighbor classifier is chosen, the trained model examines the labels of the 10 closest neighbors (Fig. 36) to the observed object. Different types, e.g., voting for counting the density, or volumetric for distance measurement, are possible. In a simple distance procedure (e.g., Euclidean Distance), the classifier compares the labels of the neighbors to determine which class the object is submitted to. In Figure 36 the new observation is assigned to the class "good" with 70% (7 good, 3 waste). A given procedure sample for a basic k-NN is presented in Pseudo-Code 4 [Alt92].

begin

1. Labelled Instances = $(\mathbf{x}_1, c_1), \ldots, (\mathbf{x}_n, c_n)$
2. New Observation $\mathbf{y} = (y_1, \ldots y_n)$
3. Instance = $(\mathbf{x}_i, c_i)$
4. Calculation of distances for each instance = $d(\mathbf{x}_i, \mathbf{y})$

   Euclidean Distance

   $$d_s(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

   Mahalanobis Distance

   $$d_{st}^2 = (\mathbf{x}_s - \mathbf{y}_t)\mathbf{C}^{-1}(\mathbf{x}_s - \mathbf{y}_t)^T$$

5. Ordering from low to high: $d(x_i) > d(x_{i+1})$ in $\mathbf{D}_y$
6. Selection of k-nearest instances to $\mathbf{y}$ in $\mathbf{D}_y$
7. Assign $\mathbf{y}$ to majority labels from selection

end

Pseudo-Code 4: Procedure of a basic k-NN classifier with exemplary distance metrics

4.2.5 Naïve Bayes

The Naïve Bayes is a cost/risk based probability classifier describing different classes by distributions [Hay98]. According to the type of distribution of the observations in each class, the probability density estimate is trained for each class in the model to be used for prediction.

4.2.6 Comparison

The previous explained methods were chosen according to their behaviour in manufacturing approaches. Support vector machines, decision trees, and k-nearest neighbor classifier methods are *State-of-the-art* components in optical control systems in polymer film industry, as displayed in Table 6. Neural networks are applied more often for high dimensional classification problems.

Table 5 Standard Methods Compared by Polymer Industry Criteria

Method: \ Criteria:	Training Effort	OCC Support	Applied in Polymer Film Industry	Other Research Applications	Speed	Memory Usage
Neural Networks	High	References	Optical Control Systems	–	Fast	High
Support Vector Machines	Low	References	Optical Control Systems	Delivery Control	Medium	High
Decision Trees	Low	–	Optical Control Systems	–	Fast	High
k-Nearest-Neighbor Classifier	Low	References	Optical Control Systems	Fibre Pattern Recognition	Slow	Low
Naïve Bayes	Low	–	–	–	Medium	Low

OCC support, applied systems and references in polymer film industry were the main criteria for method selection. A description about their industrial application and modification types will be given in Chapters 5 and 6.

4.3 Modifications of Classification Methods

Classical analytical methods for manufacturing diagnostics are bounded by sensor property and data quality problems. The implementation of classical analytical tools implicates an adaptation to dynamical manufacturing influences (e.g., transmission problems and slowly drifting deviations).
Therefore, modified machine learning methods for diagnostics achieve optimal results in anomaly and novelty detection when the limits of the operators' experience are exceeded. Their capacity of adaptation to changing problems, monitored and analyzed in real time, makes them essential for reinforcement tasks.
In the following, four *State-of-the-art* classification methods from Section 2.2 were selected by different criteria, as shown in Table 6. Applied methods in polymer film environment were mainly focused for further research. The modification of these classification methods to improve anomaly and novelty detection is discussed in this Section.

4.4 *One-Class Classification* and Trajectory Surveillance

A multi-class classification problem with a sparsely filled class of objects makes sampling for training impossible. A typical problem in the manufacturing industry occurs when many types of PLC by different vendors are installed on one machine, controlling sensor points (e.g., 200) that send different failure protocols or unknown analog signals. On the other hand, the machine condition is not well stored for production processes. New states appear that need to be classified, which cannot be simulated within real-time production. Therefore, the amount of faulty samples for problematic states needs to be equalized by cost matrices or by modified classification methods, so-called *One-Class Classification* (OCC) [[S+05a] [R+98]].

4.4.1 *One-Class* Support Vector Machines

Novelty, anomaly, and outlier detection are parts of the *One-Class-Classification* field. Methods for reconstruction, boundary, or density estimation differ in their ability to extract dataset information. An exemplary boundary method, the support vector data description by Tax (2013). [[Tax13] [D+07] [S+05b] [SV11]]. The support vector data description (svdd) is a modification of the support vector classifier by Tax (2013), with a hypersphere around the non-sparsely sampled class. The boundary is optimized with an RBF kernel and different rejection rates (requiring quadratic programming). The modified procedure of the standard support vector machines is displayed in Pseudo-Code 5.

begin

1. Define $\mathbf{X}$ as Input and $\mathbf{Y}$ as Label
2. Categorization in one/two classes $\mathbf{Y} = (1, -1)$
3. Training Target-dataset $S = ((\mathbf{x}_1, \mathbf{y}_1), ..., (\mathbf{x}_t, \mathbf{y}_1))$
4. Sparsely Outlier-dataset $S = ((\mathbf{x}_{t+1}, \mathbf{y}_{-1}), ..., (\mathbf{x}_{end-t}, \mathbf{y}_{-1}))$
5. Sample distance to hyperplane/hypersphere $d = y_i(\langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b)$
6. Boundaries/Radius around Target Datasets
7. Minimization of distance/volume (e.g., RBF kernel)
8. Compare d_i with d_{i+1}
9. Selection of maximum margin hyperplane/hypersphere d

end

Pseudo-Code 5: Procedure of support vector machines one-class classifier for hyperplane/ hypersphere calculation

The second modification is called incremental support vector machine (incsvdd), without quadratic optimization but with any possible kernel [Tax13] [D+07]. The procedure for boundary support vector one-class classification applications on sparsely sampled outliers examines distances to hyperplanes, or hypersphere, so called boundaries around the target datasets according to Tax (2013) and Schölkopf (2001). The best margin hyperplane, or hypersphere is selected. Less objects for scaling are needed compared to density estimations [S+99].

4.4.2 *One-Class* k-Nearest-Neighbor Classifier

The k-nearest-neighbor data description ($k$nndd) is an advanced method for high-dimensional feature space distance observations on neighbors [[Tax13] [D+07]]. The following Pseudo-Code 6 is accorded to Tax (2013), called nearest neighbor description.

begin

1. Labelled Target Instances = $(\mathbf{x}_1,\, c_1),\, .\, .\, .\, ,\, (\mathbf{x}_n,\, c_n)$
2. New Observation $\mathbf{y} = (y_1,\, .\, .\, .\, y_n)$
3. Calculation of distances for each instance = $d(\mathbf{x}_i,\mathbf{y})$

   Euclidean Distance

   $$d_s(\mathbf{x}_i, \mathbf{y}) = \sqrt{\sum_{j=1}^{n} (x_{ij} - y_j)^2}$$

4. Compare density of new obs. with nearest neighbor density
5. Local density equal or higher to the density of nearest neighbor?
6. Observation accepted/denied

end

Pseudo-Code 6: Procedure of a one-class nearest neighbor description with Euclidean Distance calculation

The procedure of the nearest neigbor description algorithm compares densities of neighbor objects to new observations.

4.4.3 *One-Class* Neural Networks (Auto-Encoder)

The auto-encoder neural network (autoenc dd) is a neural network trained to reconstruct the input pattern x at the output neuron(x) of the network. The difference between the input and output patterns is used as a characterization of the target class [[Tax13] [D+07] [Bis95]].

begin

1. Create architecture of neural network ($e.g.,\ 4 - 7 - 2$)
2. Randomly generate initial weights
3. Calculation of every neuron (input-hidden-output layer)
4. Calculation of difference between input and output pattern

   $f(\mathbf{x}) = (\mathbf{x} - NeurN(\mathbf{x}))^2$

5. Threshold depends on the chosen rejection ratio:

   Final Dataset = Similarity Transformation of (Trained Net, $\mathbf{x}^T$)
   Result = $\sum(\mathbf{x} - FinalDataset)^2$
   Fraction = Part of Neurons to be activated
   Threshold = (Result(Fraction) + Result(Fraction+1))/2;

6. Error signals for weight adjustment to approximate mapping

end

Pseudo-Code 7: Procedure of the auto-encoder neural networks one-class classifier according to Tax (2013), Bishop (1995)

The procedure (Pseudo-Code 7) of the auto-encoder neural networks from Tax (2013) according to Bishop (1995) uses the difference $(x - NeurN(x))^2$ between input and output pattern as a characterization for the defined class. A threshold depends on the set rejection ratio for the data.

4.4.4 NOVAS and NOVCLASS

The NOVAS method is a dedicated neural networks classifier. The associative memory systems for neural information processing (NOVAS) by König (1994) [K+94b][K+95] performs a mapping between the datasets of input and output (I^k, O^k) with a distance measure as association criteria for best matches. The distance measure in this particular case for image decision is defined by two features, the gradient direction and the magnitude, proportional to novelty, as displayed in Fig. 37, but also extendable on more features.

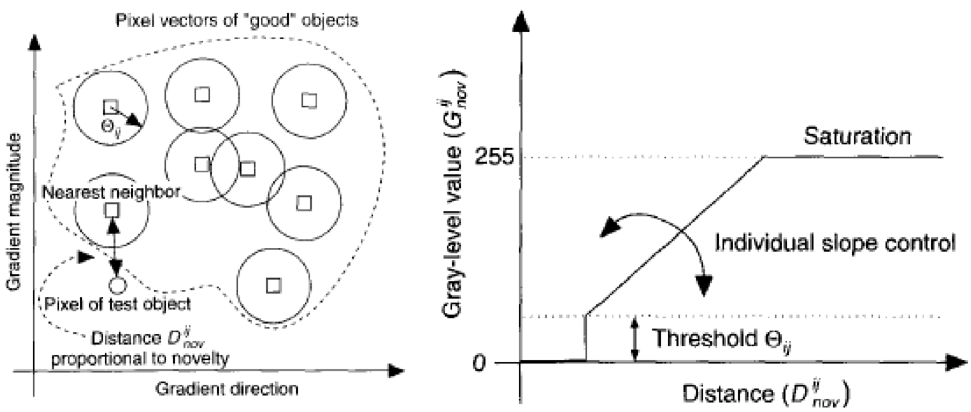


Figure 37 According to the principle of NOVAS filtering (number of features n=2) by A. König and A. Gratz (2005) [KG05]

The pixel vectors of "good" objects and "test" objects depending on the grey level value are stored in a distance matrix D^{ij}_{nov}. The following pseudo-code explains the steps for NOVAS filtering.

begin

1. n number of features, N number of good samples
2. Find the winning neuron w_{ji}^{NN} of all N good sample vectors w_{kji} that is most similar (the nearest neighbor) to current pixel vector p_{ji}

$$w_{ji}^{NN} = \min_{k=1}^{N} \left(\sum_{f=1}^{n} (p_{ijf} - w_{kijf})^2\right)$$

3. Determine the novelty distance D_{nov}^{ij} between the nearest neighbor neuron w_{ij}^{NN} and current pixel vector p_{ij}

$$D_{nov}^{ij} = \sqrt{\sum_{f=1}^{n} (p_{ijf} - w_{kijf}^{NN})^2}$$

4. Transform the distance D_{nov}^{ij} into a gray-level value G_{nov}^{ij} representing the novelty or anomaly of the current pixel

$$G_{nov}^{ij} = f(D_{nov}^{ij} - \Theta_{ij})$$

end

Pseudo-Code 8: Procedure for NOVAS filtering according to A. König and A. Gratz (2005) [KG05]

The NOVAS approach, as displayed in Pseudo-Code 8, in neural network terminology makes a minimum search in a first step, then using a nonlinear transfer function with a sensitivity threshold Θ_{ij} to calculate distances which are transformed into grey-level values. The threshold Θ_{ij} is defined as the maximum distance between two nearest neighbor objects of all sample objects. The spherical adjustment is extendable with outliers. In this case the NOVCLASS classifier, representing one channel of NOVAS for random numbered feature space, was developed by König and Gratz (2005), using hypersphere concepts for non-parametric mapping of regions [KG05].
The NOVCLASS code was implemented in Matlab for further investigation.

4.4.5 Comparison

The previous explained OCC methods were selected regarding exclusive criterias from polymer manufacturing. The criteria from Table 6 were chosen according to the OCC methods behaviour in polymer manufacturing problems. Important issues are training behaviour, understanding the methods structure, references, recommendations, and their degree of popularity.

4.5 Novel Adapted Approach

A novelty recognition approach improves the detection of abnormal behavior in new datasets if only one class of extensive knowledge is available. The methods used in this case could be non-parametrical and could describe a complex multispherical structure in several dimensions (radial centers due to the dataset distribution). In a first, e.g., randomized hold-out selection, the data could be divided into a training set and the rest, followed by a second hold-out selection into the validation and testing datasets, afterwards implemented into a sequence as exemplary shown in Fig. 38.

Table 6 Compared OCC methods by different criteria

Method: / Criteria:	Easy to train	Easy to understand	Reference Projects	Recommended	Degree of Popularity
Support Vector Machines	Yes	Yes	No	Yes	High
Neural Networks (Auto-Encoder)	No	No	No	Yes	High
k-Nearest Neighbor Classifier	Yes	Yes	No	Yes	High
NOVCLASS	Yes	Yes	Yes (Industry & Research)	Yes	Low

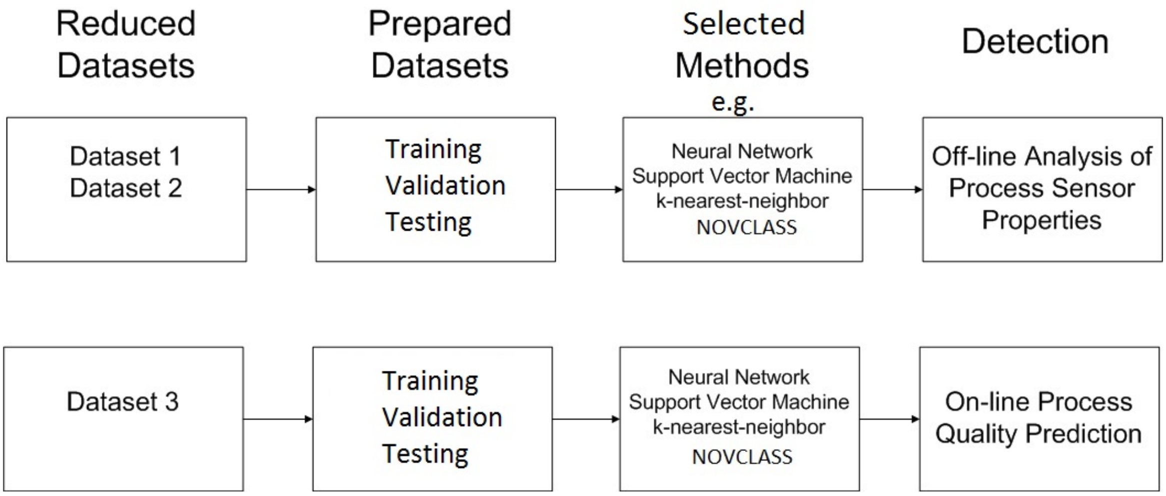


Figure 38 Cycle of dataset selection, preparation, classification method, and implemented detection system

The three queried datasets (Datasets 1, 2, and 3) from the database will be reduced from higher dimensions to lower process attributes. The next part of the setup shall train, validate and test normal and faulty operating conditions, by different OCC classifier types for recognition, concluding in an off-line analysis of the process sensor properties and an on-line process state prediction (Fig. 38).

4.5.1 Methods Evaluation

Following the Chapter 2, now discussed methods, support vector machines, k-nearest-neighbor classifier, NOVCLASS, and auto-encoder neural networks shall be applied, with possible modifications as listed in Table 7, according to König et al. (1994, 2005) and Duin et al. (1998) [[K$^+$94a] [K$^+$94b] [K$^+$95] [KG05] [D$^+$07] [Tax13] [M$^+$01]]. The exemplary parameter range of preliminary adjustment for OCC-tested methods is shown in Table 7. Within these parameter ranges the testings are possible. E.g., the chosen number of hidden neurons (1–200) with rejection rates from 0.00000001 to 1 for the neural networks (auto-encoder) could be trained to reconstruct the input pattern x to the output neuron(x); the support vector machines (incremental) with kernel for sigma 1–20 and gamma 0.001–0.1 and the k-nearest-neighbor classifer (data description)

Table 7 Parameter ranges of preliminary tested methods for OCC

Methods	Classifier Type	Settings
Neural Networks (Auto-Encoder)	OCC	HiddenNeurons= from 1 to 200; Rej (Rejection) = from 0.00000001 to 1
Support Vector Machines (RBF)	OCC	Sigma= from 1 to 20; Gamma= from 0.001 to 0.1; Rej= from 0.00000001 to 1
k-Nearest-Neighbor Classifier	OCC	k= from 1 to 30; Rej= from 0.00000001 to 1
NOVCLASS	OCC	ScaleFactor= from 0.1 to 2

with neighbors 1–30 shall be modified with, e.g., the Data Description Toolbox 2.0.0 and the Pattern Recognition Toolbox 5.0 [[D$^+$07] [Tax13]]. Another modified classifier for OCC shall be tested with the NOVCLASS method in Matlab. By changing the settings for each of the exemplary methods on different datasets, the best accuracy shall be searched for, followed by a significance test (e.g., t-test). All methods train distance ranges for datasets. The positively or negatively classified datasets and their corresponding accuracies depending on the optimized boundary settings realized in Matlab R2012b (8.0.0.73) will be based on trained distance ranges [The15].

5 Trajectory Process Visualization

Today's production machines start capturing datasets from all factory points, and the upstream process steps need to be more adaptable to incoming reports on deviations. The generic way of investigating different process conditions shall be upgraded by a sequential event analysis approach for earlier reaction, as a proprietary contribution to the previous OCC results. One upstream condition shall be replaced by more conditions in sequence, to give a precise prediction of upcoming events. In the following, the extension from the classification of an individual process state per time step to the classification of a temporal sequence or chain of process states, denoted as trajectory, will be pursued.
Timely recognition of trajectory datasets from different process conditions, combined with *State-of-the-art* classification methods, allows faster reaction within the first production minutes in a generic but, for the polymer film industry, novel way. The methodological approach from the audio signal-converting industry helps to understand sequential events for the investigation of one particular polymer production process. In Fig. 39, an exemplary trajectory from the polymer process (extrusion line) is displayed.

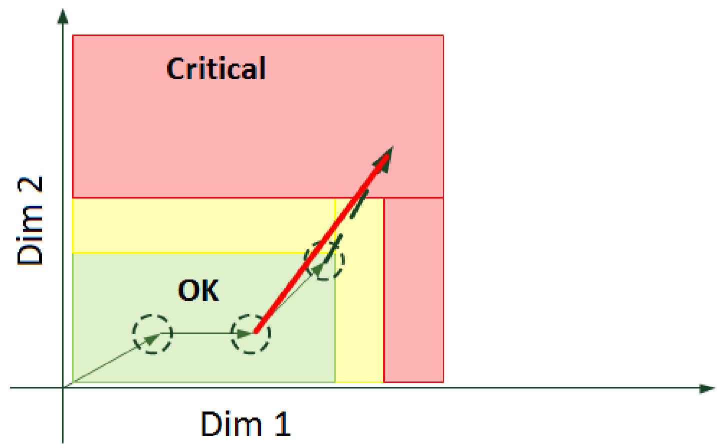


Figure 39 Process trajectory from extrusion process (pressure) with abnormal behaviour

The time shift from the second to the fourth state implies about 10 states in process surveillance. Fig. 39 gives an impression of the average trajectory behavior. Section 5.1 starts with the trajectory field of applications, continued in Section 5.2 with a novel modification for manufacturing integration. The dataset acquisition and the use of mutual information for finding the best sliding sequences are also described in Section 5.2, followed by a description of the general methods.

5.1 Field of Applications

The trajectory investigation mostly focuses on, e.g., the search for causes of water or air pollution (back trajectories) by exploring meteorological patterns in emissions, health monitoring for disease trend recognition, aircraft behavior monitoring to ensure safety in the air space, packet transmission in dense network types for faster paths, or trajectory control optimization of industrial robots for raising productivity. Different industry types from the medical industry to the automobile industry or the meteorological services use trajectory functionalities to recognize patterns and for simulation [[W+94] [HvS00] [L+08] [V+02]]. Novel trends look at stock exchange markets, their trajectory behaviors in relation to market changes are promising, or yield optimization approaches in microelectronics, and their adaptation to broader industrial fields. The industrial approach started with computer-aided robotics and trajectory controllers. Optimization of these systems led to lower energy consumption, higher yield due to faster output, as , waste reduction, higher safety issues, and cost reduction regarding human resources [[M+01] [KG05]]. Still there is a lack of having research approaches for such trajectory developments in the particular field of polymer film production. Manufacturing machine operators make high demands on time series monitoring, and attribute behaviour over long periods. The *Trajectory-Visualization* for advanced monitoring and recommendation purposes offers improvements to successfully achieve such approaches for yield optimization.

5.1.1 Trajectory Visualization

A trajectory visualization of condition attributes (e.g., pressure, temperature) in 2-dimensional space could base on a selection of process condition attributes. Each trajectory additionally consists of different timestamps due to abnormal processing behavior. The production of one product takes about 110 min, but is running 24 hours/7 days for the complete production line. Therefore, each trajectory might consist of 10–110 timestamps or more. The usual length depends on material quality and process conditions. In Fig. 40, multiple timestamps from raw extrusion data (pressure) are displayed. The red lines represent the abnormal process conditions, and the green lines, the normal ones.

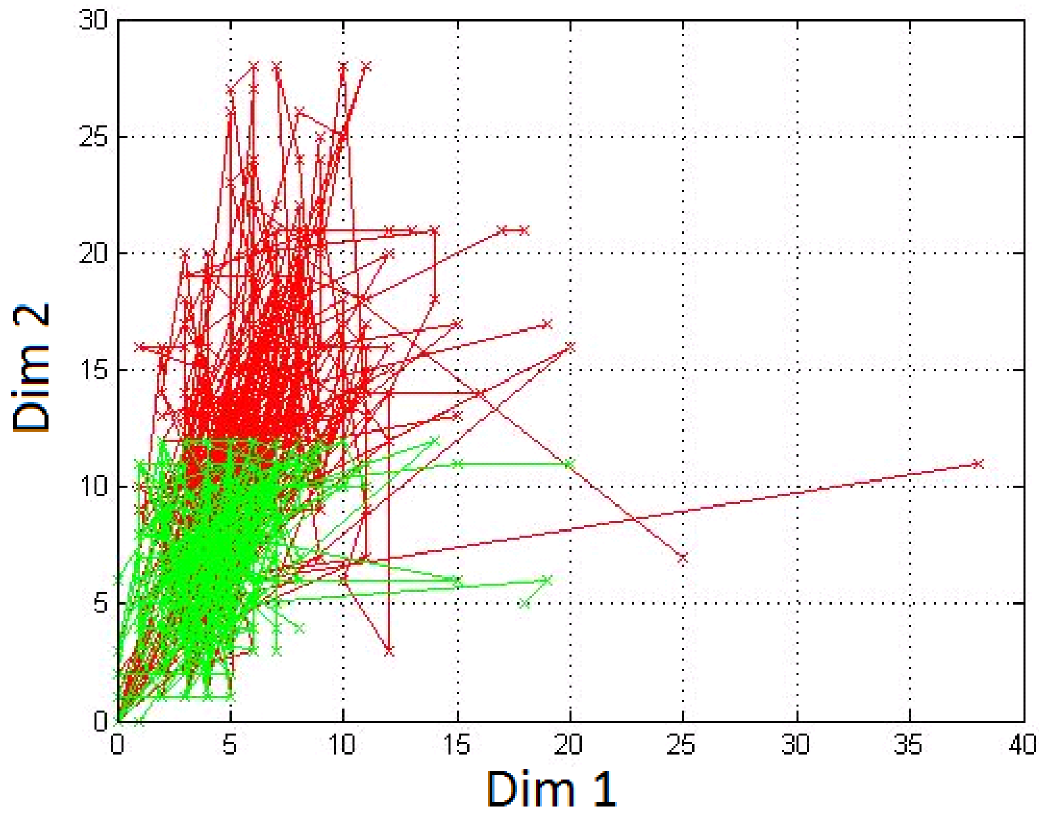


Figure 40 257 trajectories in a coordinate system for 2 specified dimensions from extrusion raw data

The trajectory datasets will be projected in a Cartesian coordinate system and could be reduced from, e.g., 13 timestamps to the same size of 3 timestamps, with mutual information feature selection in Matlab and STATISTICA ($\tilde{\chi}^2$, p value). This way, all trajectories will have the same scanning window, which is displayed as the sequence of events as shown in Figure 40. Examinations of the trajectory high-dimensional space offer the possibility for feature selection, to choose the best visualization attributes. Each timestamp consisting of multiple attributes; each trajectory comprises of same timestamps after standardization. Thus, more than 60 attributes could be distinguished for classification.

5.2 Novel adapted Approach

The predefined OCC classification results of the abnormal and normal process conditions from the real production datasets were trained, validated and tested with *State-of-the-art* classification methods in a second step. Exemplary time series taken from real-time production ranging from 10 min to several hours have to be standardized and modeled, as shown in Pseudo-Code 9.

begin

1. Define $\mathbf{X}$ as input and $\mathbf{Y}$ as Label
2. Categorization in two classes $\mathbf{Y} = (1, -1)$
3. Training dataset $S = ((\mathbf{x}_1, \mathbf{y}_1), ..., (\mathbf{x}_t, \mathbf{y}_t)) \subseteq (\mathbf{X} \times \mathbf{Y})^T$
4. Sample distance to hyperplane $d = y_i(\langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b)$
5. Compare d_i with d_{i+1}
6. Selection of maximum margin hyperplane d
7. Repeat until best result found

end

Pseudo-Code 9: Procedure of the used *State-of-the-art* support vector machines classifier for trajectory visualization

The procedure for *Trajectory-Visualization* with the *State-of-the-art* support vector machines is explained in the Pseudo-Code above. The input dataset $\mathbf{X}$ was categorized in two classes Y. By hold-out selection the training set was separated, and in a following hold-out selection, the evaluation and testing datasets were defined. For new observations the sample distance to the hyperplane was calculated and compared. After the selection of the maximum margin hyperplane the procedure was repeated until best result could not be optimized anymore.

5.2.1 Methods Evaluation

The following methods, listed in Table 2, were applied by adjustments within ranges: support vector machines, k-nearest-neighbor classifier, naïve Bayes, neural networks, automated network search (MLP), and boosted trees. These *State-of-the-art* multi-class classifiers are based on predefined knowledge of nearly all possibly occurring cases.

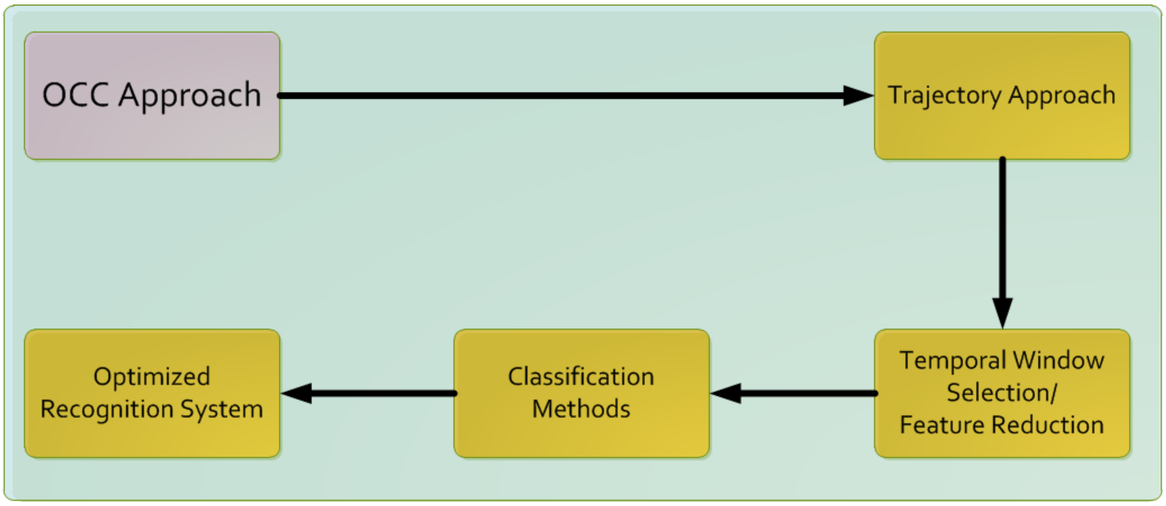


Figure 41 Cycle of the OCC approach to a trajectories approach, feature reduction, classification approach and optimized system

The cycle of a possible OCC approach, combined with a trajectory approach, feature reduction, classifier investigation for optimized recognition is displayed in Fig. 41. In Table 8, the parameter ranges of the selected multi-class methods are displayed. All methods will be used for training, validation, and testing by hold-out selection. The randomized selection divides the datasets into training, validation, and testing sets. Different distributions will be tested.

Table 8 Parameter ranges of selected multi-class methods for trajectory classification

Methods	Classifier Type	Parameter Ranges
Neural Networks	RBF	HiddenNeurons= from 1 to 200; Rej (Rejection) = from 0.00000001 to 1
Support Vector Machines	RBF	Sigma= from 1 to 20; Gamma= from 0.001 to 0.1; Rej= from 0.00000001 to 1
k-Nearest-Neighbor Classifier	–	k= from 1 to 30; Rej= from 0.00000001 to 1
Naïve Bayes	–	Prior probability; Threshold = 0.0001; Seed from 1 to 1000
Automated Network Search	MLP	Hidden units from 1 to 50; Networks to train from 1 to 200; Networks to retain from 1 to 50; Error function = sum of squares, cross entropy; Activation function (hidden, output) = Identity, Logistic, Tanh, Exponential, Sine; No weight decay; Seed from 1 to 1000
Boosted Trees	–	Missclassification cost = equal; Prior probabilities = estimated, equal; Learning rate = from 500 to 2000; Number of additive terms from 100 to 300; Random test data proportion from 0.1 to 0.9; Subsample proportion from 0.1 to 0.9; Seed from 1 to 5; Minimum n in child node = 1; Minimum n of cases = 40

The possible parameter range for the generally used methods is reported in Table 8. The approach will be done with different multi-class methods on different datasets. The best accuracy will be searched for by changing the settings for each method followed by a significance test (t-test).
For the experimental settings, the neural networks with radial-basis-function could be used to reconstruct the input pattern x to the output neuron(x) of the network with 1–200 tested hidden neurons. The incremental support vector machines works with any kernel for sigma 1–20 and gamma 0.001–0.1. The k-nearest-neighbor classifier, the automated network search (MLP) and a boosted tree method are included too. By changing the settings for each of the displayed methods on the datasets, the best accuracy could be searched for, followed by a significance test (t-test). Their accuracy depends on the optimized boundary settings to be realized in Matlab R2014b (8.0.0.73) and STATISTICA 12. The experiments and results are explained in the Chapters 7, 8, and 9.

6 The System Architecture

In this Chapter, the system architecture is explained from the conceptual and from the currently implemented point of view, for the data processing approach to the polymer film manufacturing case study. Inspirations are taken from approaches and architectures of microelectronics manufacturing processes [[M+01], [KG05]]. The level of automation is regarded from a holistic viewpoint and from a detailed perspective, with regard to [[K+13] [K+95] [K+94b]].

6.1 The Conceptual Design

The holistic overview concept considers “Industrie 4.0” as the main objective that has to be reached. This includes the Digital Factory and Smart Production, which is totally cross-linked, thus integrating customers and suppliers along the value-added chain, by using big data methods for data overload, *Self-X, Self-* Properties (abbreviated as Self-X in the following)*, and approaches for *Self-Optimization*, *Self-Configuration*, *Self-Organization*, *Self-Protection*, *Self-Healing*, *Self-Explaining*, and *Context Awareness* in *Cyber-Physical-Production-Systems*, and *Human-Machine-Interactions* as operators support.
The focus on advanced manufacturing processes becomes more important in times of individualized mass customization and product life cycle management with regard to predictive analytics, cyber-security, and clouds. Research advancements such as *SElekt I4.0* by the BMBF (2014) focus on these fields to support approaches of that kind for *Industrie 4.0* [[BMB14] [Win13] [Bun15]].
The conceptual proprietary design for the polymer film industry was developed in three graded layers. Each layer describes a further advancement of the previous one, building on each other.

The Layers of Automation

The three layers of automation viewed from the polymer film industry case study are presented in the following Fig. 42. The layers are divided into the “’*Acquisition Layer*” (A), the “*Monitoring & Recommendation Layer*” (B), and the “*Adaptation Layer*” (C).

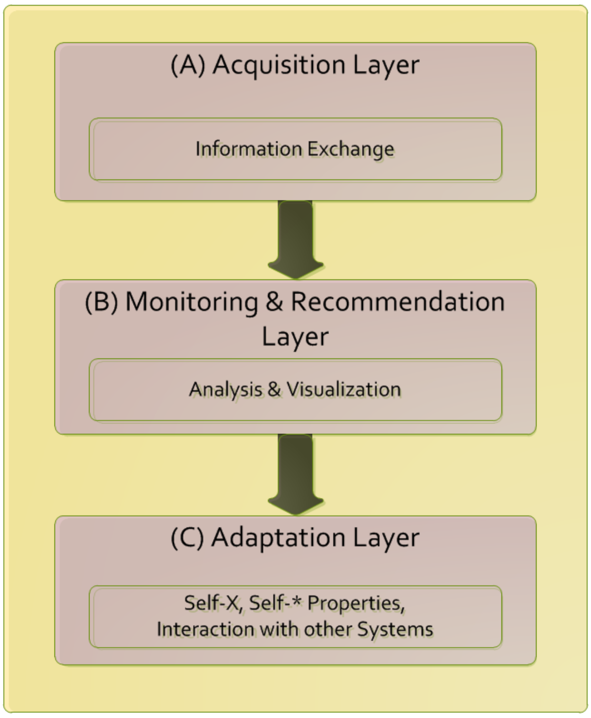


Figure 42 Framework for autonomic development from the polymer film industry view: (A) Acquisition Layer, (B) Monitoring & Recommendation Layer, (C) Adaptation Level

Starting with the "*Acquisition Layer*" **(A)**, the machines[22] as minor objects are independent working elements. Each object is managed by a supervisor who makes manual adjustments. The objects are linked to a network, such as TCP/IP, providing datasets to a centralized database for consolidation. The "*Monitoring & Recommendation Layer*" **(B)** systems offer technologies for pattern recognition, prediction issues, and for recommendations to the machine operators or to the supervisors for process improvements and advanced manual control. In the "*Adaptation Layer*" **(C)**, the machines automatically take actions to adjust the machine settings for process improvements according to the given information. The manual intervention by experts has been abandoned. The machine is guided by objectives and interacts with other systems such as, e.g., automated guided vehicle systems for picking up finished products.
In this Section, the categorized layers (A), (B), and (C), building on each other beginning with (A), were investigated from the *Polymer Industrie 4.0* design view.
The following, more detailed comprehensive report discusses the categorized automation layers, adapted to the polymer film case study from Chapter 3, regarding the future possible *State-of-the-art* compared to the current implementations as described in Section 6.2.

6.1.1 The Acquisition Layer

In the **Acquisition Layer**, the minor objects (i.e. the machines) are independently managed by supervisors. The process lines consist of two to three consecutive machines, the sub-divisions include connected process lines. The divisions consist of several machines, process lines, and sub-divisions connected within a product life cycle process.

[22] Workstation, Chapter 3

The plants comprise all divisions and the markets connected to the divisions, and the corporate groups include several plants each Fig. 43.

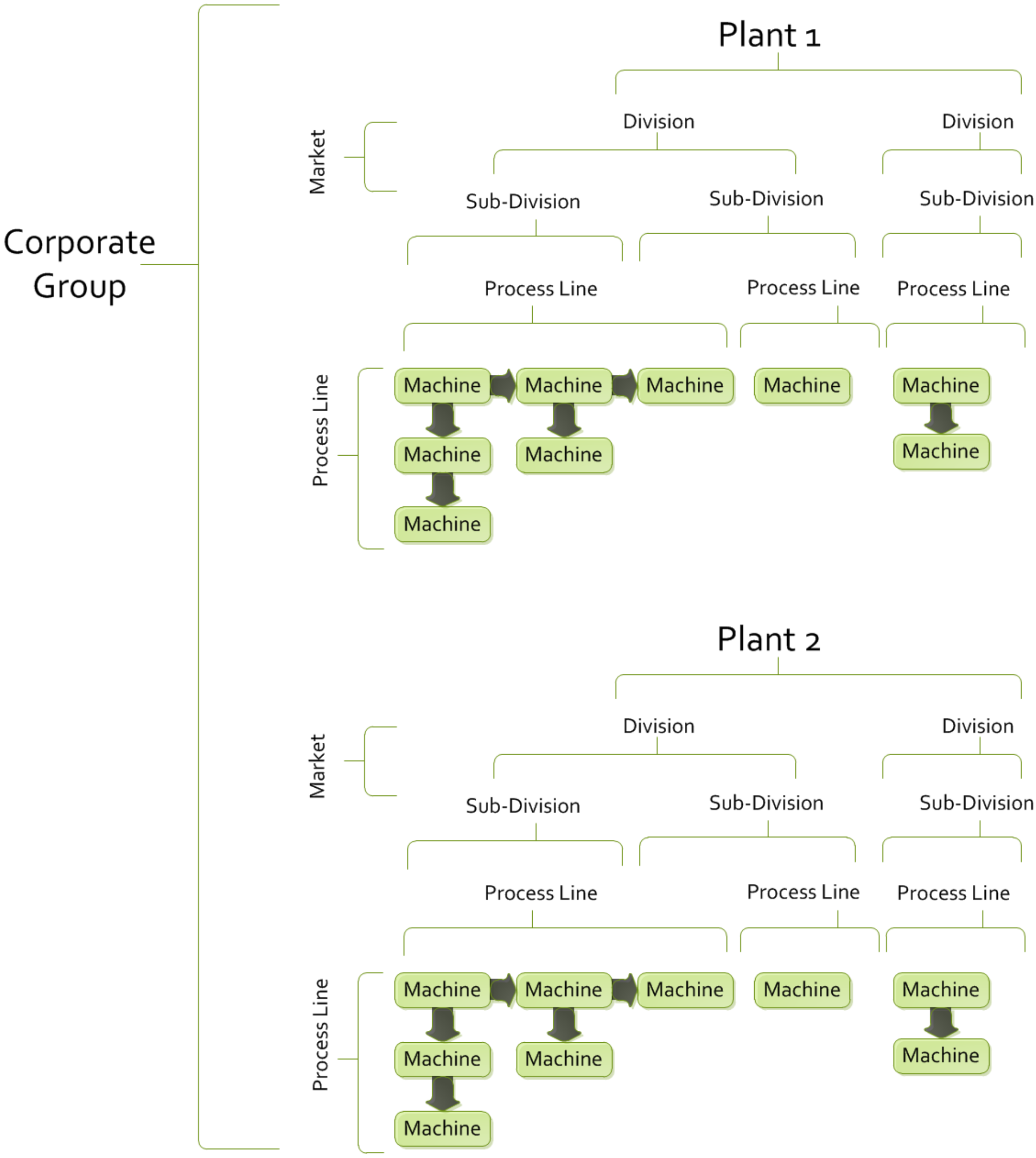


Figure 43 Objects from a corporate group and market view

The locally acquired data at the machine (minor object) represent a bottom-up view, to be used in higher aggregations. The object's settings, which are locally stored, are adjusted by technical and process experts, if production problems occur.

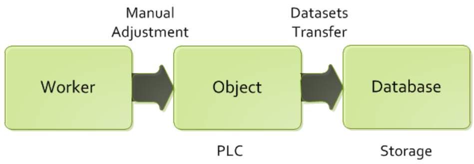


Figure 44 Single object and database network connection

Simple handling is supported by technical improvements concerning partial local sensor monitoring and database storage (Fig. 44). This layer realizes the first part of the *Smart Production and Digital Factory*, which illustrates the wirelessly cross-linked (by the *IT-Infrastructure*) machine network, with sub-divisions of several connected machines in a process line, consolidated to divisions within a plant, in a corporate group producing *Just-In-Time* products for the market. This so-called *Cyber-Physical-Production-System* transfers cloud information (such as instruction or setting changes) between decentralized intelligent units, e.g., in an embedded sensor system with digital memory, or to a centralized data storage unit (Fig. 45).

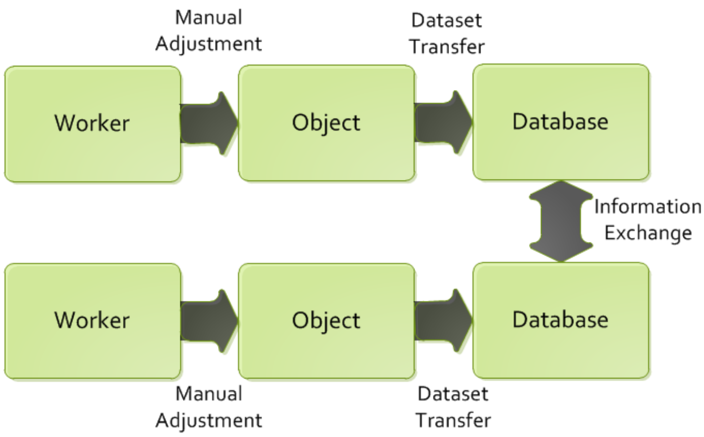


Figure 45 Multiple object and database network connection

For this reason, an established network with memory objects within the IT-Infrastructure is required to handle high data volumes. Data sources, such as the In-Memory technology, NoSQL, or Cloud, and their fast 1-GB wireless link manage the information flow between and within objects. Supported by database technology for data transfer, the *Smart Factory* accesses technology platforms and capacity trading centers to use market information as machine-machine or machine-supplier support. With regard to the polymer film case study, from a futuristic point of view, the following scenario can be envisaged:
The overall machinery consists of 65 high-complexity processing lines producing up to 1200 kg product/h, located in 8 interlaced divisions. Sensor points acquire real-time data from the production process, which is transferred to databases by data collectors. Each of the four to five machine PLCs, e.g., SoftSPS or S7, is permanently wirelessly connected to the machine network VLAN, communicating in milliseconds with different decentralized process databases for data storage. Several machines are linked by the product, e.g., RFID, or virtual identifiers, over two to four sub-divisions within a *Cyber-Physical-Production-System*, and to suppliers and customers for product identification and traceability, exchanging more than 30,000 process quality datasets per minute, as

well as deviations and recommendations in both directions; between database systems, considering growing data volume and improving data quality.
Each machine or sub-division is linked to its production division, getting additional data from Engineering, Quality Management, Human Resources, Logistics, Supply Chain, Controlling, Distribution, and External Sources (e.g., Weather Forecast, Social Media). Many decentralized parallel-processing outsourced storage systems for high-performance computing with Big Data (e.g., Hadoop, NoSQL, or In-Memory technology) are connected to a decentralized flexible data warehouse cloud, providing prepared datasets in real time for further processing in, e.g., *Computer Assisted Quality Control* (CAQ) enterprise management systems or sequence planning tools using smart technologies.

6.1.2 The Monitoring & Recommendation Layer

The **Monitoring & Recommendation Layer** is built on the *Acquisition Layer*; statistical methods are used to analyze the information transfer within such systems. Today's Big Data methods employ supervised learning algorithms comprising neural networks, auto-encoder neural networks, support vector machines, decision trees, boosted trees, k-nearest-neighbor, naive Bayes, NOVAS, and NOVCLASS (see Section 4), and modifications of these methods to handle underdetermined observation classes can be implemented in a multitude of application types for specific tasks. The objects (e.g., machines) use variable methods for intelligent behavior in order to detect unknown conditions, sensor failures, or energy consumption. The digital product identification allows the transparent traceability of objects through the supply chain. Pattern recognition systems for similar products support the planning department, and simulations of alternative processing routines within the company structure offer new advantages, by suggesting additional or different technologies for manufacturing, to raise the efficiency and to help prevent waste. Cause Sourcing, Product Identification and Traceability, and Simulations provide complex process knowledge for recommending new machine settings by *Assistency Systems*, e.g., for lowering energy consumption or solving actuator deviation problems. The integration of customers and suppliers is of high importance due to the possibilities of earlier process interventions and faster adaptation of the sub-divisions to manufacturing condition changes. Alternative routes, the real-time prevention of waste, predictive maintenance, and improved resource allocations are achievable.
With these targets in mind, different existing supervised, unsupervised, and modified methods explore defined datasets and prioritize attributes by feature selection. Extensive statistical models, provided by software tools such as SAS Visual Analytics, STATISTICA, Matlab & Simulink, Rapidminer, and QuickCog, need to be trained, evaluated and tested with datasets collected in the last 10 years from objects such as machines, sub-divisions, divisions, plants, markets, and corporate groups; the models are cyclically compared and visualized graphically. A multitude of novel approaches such as modified recognition methods or visualization techniques are automatically integrated into the loop to find the best prediction for each minute, hour, and day, to support yield improvements.
The prediction error of unknown conditions needs to be totally minimized, and the prediction time point would be nice to be set as early as possible, e.g., 1–2 days in advance for machines and sub-divisions, 3–4 months in advance for divisions, and 1–2 years in advance for plants, corporate groups, and markets.
The extracted object results regarding accuracy, cost, waste, and time and their settings are imported into a calculation system for manual and automated simulation of the

best processing conditions, because of unknown datasets that need to be analyzed in advance of occurring events for each regarded object, e.g., division losses due to machine shutdown or the probability of an energy net stability fee versus a renewable energy law fee. The identified pattern from all acquired objects is translated into performance measures for fast interpretation and downstream monitoring. The results gained from the interactive analysis and knowledge acquisition are visualized by a monitoring system (Fig. 46), capturing performance measures (e.g., deviations and cost) from the above-mentioned objects in real time.

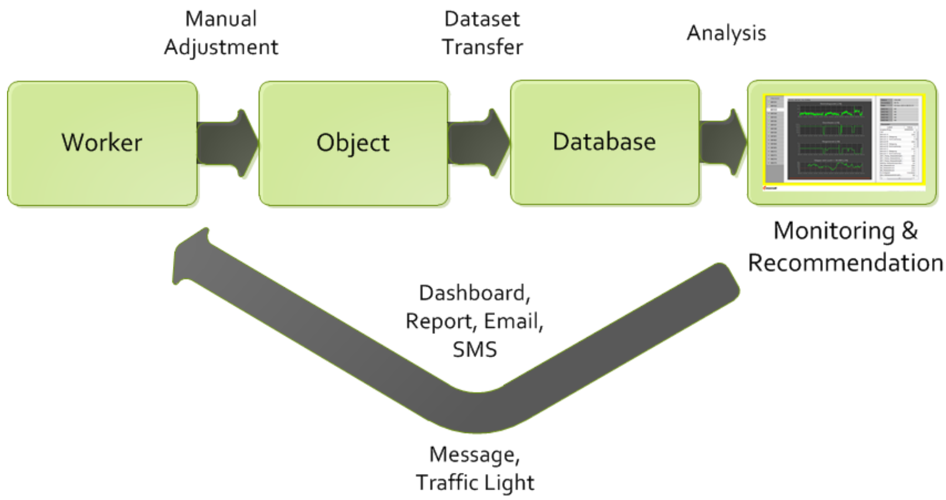


Figure 46 Real-time computer-aided manufacturing processing action & recommendation system

The *Human-Machine-Interface* idea yields opportunities for recommendations by informing employees via *IT-Infrastructure* systems. Dynamic and static statistical limits are established for all acquired data performance types in the data warehouse system, reduced to each single attribute. The structure is related to traffic light systems for easy understanding, giving status information on objects, groups of objects, and their predicted behavior in advance, realized in three necessary types: (1) **Visualized Violation**, (2) **Recommendation System**, (3) **Silent System**.
The *Assistance Systems* offer information via webpages or standalone applications, or integrated into (e.g., via the outlook surface) status details through a dashboard functionality in different possible views. A simulation part for product calculations is needed to perform additional interactive analyses, such as price negotiations or parameter settings.
The recommendation part gives advice via interface or mobile phone, by traffic light or sound, to provide information changes to the responsible person or to appended control systems (e.g., e-mail, SMS, file transfer).

6.1.3 The Adaptation Layer

The previously mentioned monitoring and recommendation part offers information input for the following adaptive systems. If limit violations occur, an autonomous object, e.g., a *Self-Configurating* object, takes on tasks such as changing the actuator settings for deviation ratio reduction (Fig. 47).

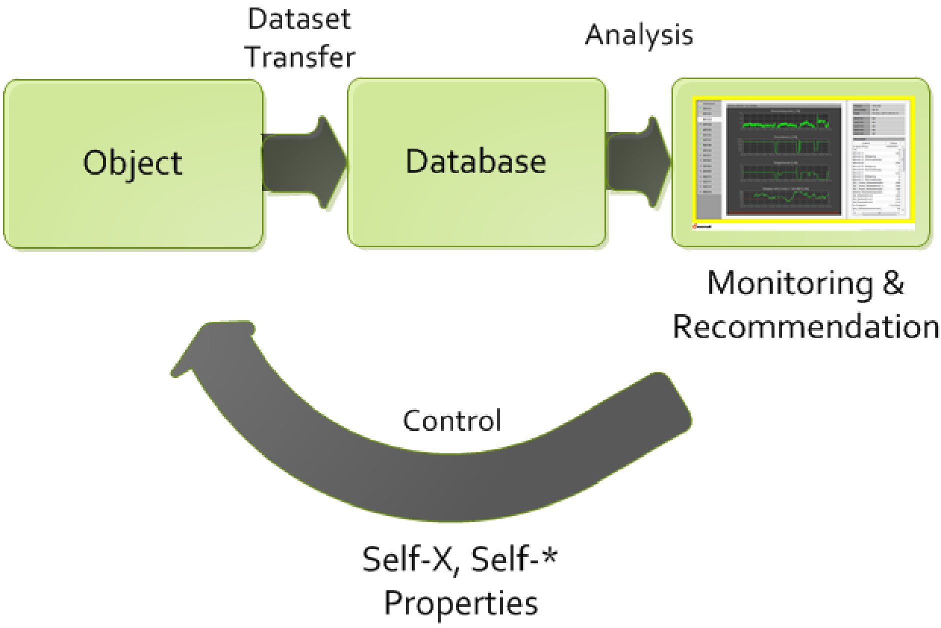


Figure 47 The object generates datasets, monitored by recommendation systems to take actions if necessary

The **Adaptation Layer** describes the autonomous behavior of intelligent objects, such as a *Self-Organizing* production capacity control for the allocation of human resources, a *Self-Optimizing* production control using intelligent sensors for releasing operators from routine tasks by *Human-Machine-Interaction*, or a tool management system within a cloud offering actuator self-optimized settings. Therefore, *Self-Monitoring* of sensor points over time and their *State-of-the-art* correlation to other sensors for diagnostic mining analysis of actuators is indispensable for failures (such as erosion, contamination, cloaking, or poisoning) to be detected by *Self-Healing* and solved by *Self-Calibration* or *Self-Trimming*, which will be investigated in future approaches in addition to this work. *Self-Monitoring* can be used for machine parts running out of tolerance within a specified time range. Early prediction avoids material waste and machine downtimes, and thus their expensive impact of several thousand euros. The integration of machine attributes into such a monitoring cycle for pattern analysis could prevent these situations and would be a profitable adaptation to the *Monitoring & Recommendation Layer* within the polymer case study.
The *Self-Configuration* of sensor points to influence the process conditions provides the machines with the autonomous control necessary to handle or prevent production problems within a quality control loop.
For instance, an automated guided vehicle system could be sent for and directed to the nearest drop station to pick up 3000 kg of waste. In another case, the preventive maintenance system could be instructed to perform electrical service at line 150 in advance of the main engine's breakdown, or to relocate the newest product type from Division 2 to the Hungarian plant for cost optimization due to new market prices. Automated guided vehicle systems are integrated into the dynamic production data cycle through an intelligent *Self-Optimizing* master control station, transporting products from portals to the *Full-Automated Self-Storage High-Bay-Warehouse*. Local autonomous objects with digital memory are able to react flexibly and use their resources more efficiently for yield optimization, which may be necessary due to changes in the product mix or because of external influences. *Self-X* in this context describes the

Self-Organization, Self-Configuration, Self-Optimization, Self-Healing, Self-Protection, Self-Explaining and Context-Awareness of autonomous embedded systems or of control stations for cable or wireless Ethernet machine-machine interactions, e.g., to choose the right machine settings to achieve a reduction in energy consumption [[W+08] [BMB14]]. All processing and workflow steps concerning globalized objectives such as material waste reduction, energy efficiency, or cost minimization are synchronized within the autonomic environment. Beside the typical control system for object adaptation, an overall system for object arrangement regarding holistic adjustment is established. It is described as a self-managing system defined by *Self-Configuration, Self-Optimization, Self-Healing, and Self-Protection.* The complete setting is globally aligned; therefore, local minima (e.g., yield losses for objects) are accepted in view of the final target of yield maximization, e.g., for the corporate group. Additionally, each control system is itself able to take actions, like questioning the applied methods and making independent replacements or calibrations (Fig. 48) [[KC03] [P+03]].

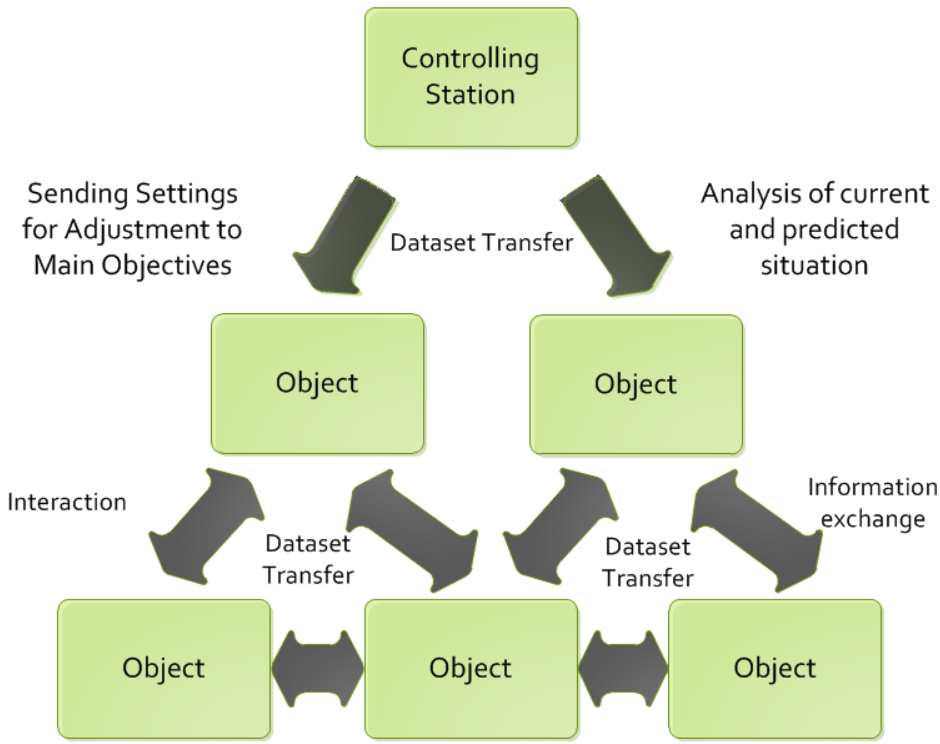


Figure 48 All system operations and object interactions are governed by business objectives

To ensure a persistent overview of the past 10 years, the present year, and the predicted 2 years, the development of object behavior monitoring in real time and, based on this, an automated report system are necessary surveillance parts. The responsible person receives information by on-line dashboards, interfaces, web pages with refreshing rates of 1 min on the computer, mobile phone, or data glasses from all objects. Typical reporting types are ratio comparisons of days, months, last years, previous objects, and other objects, as a benchmark. Therefore, it is important to know the current *State-of-the-art* and future targets for products, machines, sub-divisions, divisions, plants, corporate groups, and markets (e.g., the stock exchange market). Additional information, e.g., regarding sales quantity trends, supports the forecasts on customer and market behavior. Besides the yields, ratios, deviations and distributions, visualizations (e.g., by traffic lights) at the machine interface have a high impact by supporting operators and managers. However, the manager or supervisor is only an observer whereas the action is taken on by the object itself. A master control station, located in the cloud, assigns tasks to other local control stations for directing objects.

6.1.4 The Proposed System Architecture for Polymer Film Industry

The conceptual overview with regard to the previously explained particular layers (A)–(C) is applicable to the mentioned objects, i.e. machines, process lines, sub-divisions, divisions, plants, the corporate group, and the market, from an detached point of view, as shown in Fig. 49.

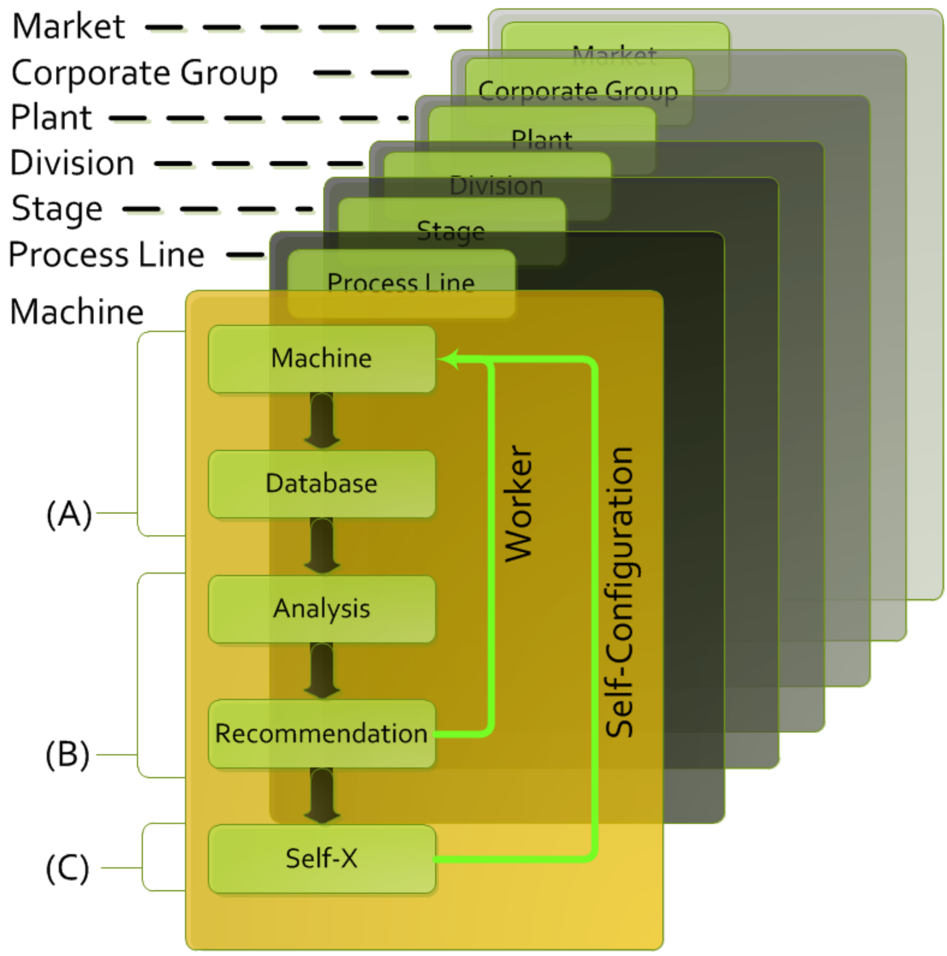


Figure 49 Proposed multi-dimensional system architecture with equal processing cycle for different objects, as machine, process line, sub-division, division, plant, corporate, and market for the layers (A) Acquisiton, (B) Monitoring & Recommendation, (C) Adaptation

The front view for the object (the machine) shows the automation layers (A) Machine & Database, (B) Analysis & Recommendation for the operator to manually adjust the Machine, and (C) *Self-X*, adjusting the Machine by *Self-Configuration*. The grey-shaded parts are copies of the front view, but replace the "Machine" by the other types of objects previously described in Subsection 6.1.1.
The proposed system architecture for one machine is applicable to several others of the

above elements from a multi-dimensional point of view, regarding novel development approaches in the manufacturing industries. Each object consists of particular attributes to be trained on a specified problem with modified methods. The system structure differs between all objects depending on the approach pursued. In the following Section 6.2, the proposed system architecture for the polymer film industry was embedded as a Real-Time Computer-Aided Manufacturing Processing Action & Recommendation System within the current implementation, according to [K^+13].
IBM (2002) conceived five levels of automation: basic, managed, predictive, adaptive, and autonomic, as a framework for autonomic computing [M. 02]. In the following, the previously developed conceptual polymer film layer view is presented, from the currently developed implementation point of view.

6.2 Current Implementation

The conceptual design was drawn in a perfect environment, disregarding industrial restrictions such as economy, and with the best available technology in the polymer film industry. In the following Section, the conceptual design is applied to subfields of the polymer film case study considering all the possible layers (*Acquisition Layer, Monitoring & Recommendation Layer, Adaptation Layer*), then implemented and subsequently planned (Fig. 50).

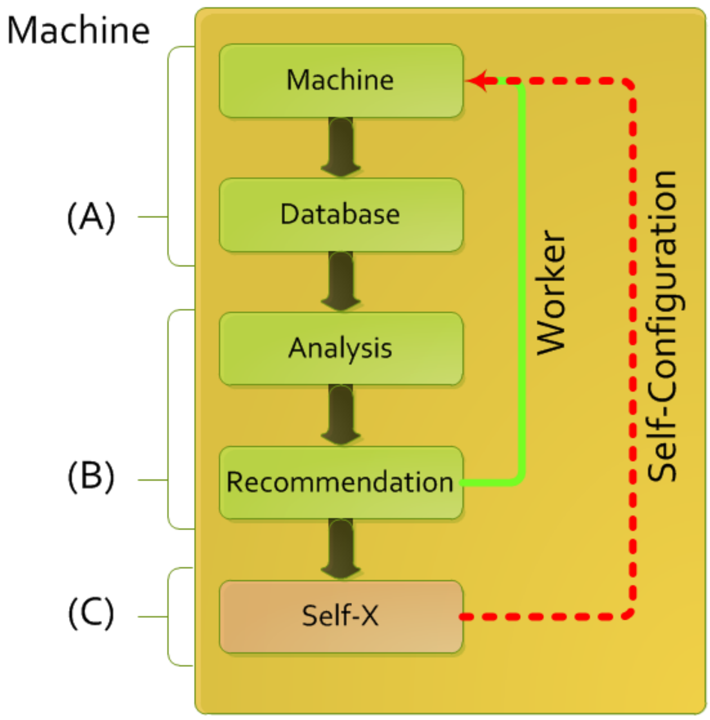


Figure 50 Proposed conceptual implementation for the case study investigation: *Machine, Database, Analysis, Recommendation*, marked green in the automation framework of the polymer film case study

The layers **(A)** and **(B)**, marked in green, could be implemented whereas the red part **(C)** needs more investigation. In the following, the possible implementation of the selected parts from Section 6.1 is demonstrated, concluding in Fig. 51. Section 6.3 gives an overview of a proposed system architecture for a Real-Time Computer-Aided

Manufacturing Processing Action & Recommendation System for the daily working routine.
Inspired by (König and Gratz, 2005) [KG05] [K+95], advanced methods in the field of polymer production process data analysis for feature reduction, novelty recognition, and interactive visualization will be investigated. The complete proposed system architecture explains the conceptual exploratory Analysis & Recommendation design of the cycle from the polymer production process to Oracle database storage, to interactive exploratory Analysis & Knowledge Acquisition, to OCC Novelty Detection, to *Trajectory-Visualization*, and finally to a process control interface with a recommendation system. Based on the referenced system architecture [KG05], the methodological approach for the polymer industry will be applied aiming at an early intervention during the production process before the occurrence of problems. The currently implemented layers are explained in the following.

6.2.1 The Current Acquisition Layer

The eight separate, highly complex machines are linked to a centralized Oracle database. The generated data is stored locally at the machine and acquired by several decentralized database systems (such as MES) and quality measurement systems, which forward selected datasets to the Oracle database. The machines are independently managed by both operators and supervisors. Their tasks include manufacturing instruction entries, adjustment, maintenance, and monitoring.
The current **Acquisition Layer** for *Smart Production and Digital Factory* was investigated in the case study with a focus on *CPPS*, their real-time data acquisition, and the complexity of cross-linked information transfer. The *IT-Infrastructure* with regard to the *KDP* from Chapter 2, the sensor data acquisition from several machines with diverse software tools, and the network data transfer were implemented within the time range from 2007 to 2015, based on the *CPPS* within the polymer film case study from Chapter 3. Comparing the conceptual design with the polymer film case from Chapters 2 and 3, similar combinations such as TCP/IP network technology, SoftSPS PLC acquisition, and local high-performance In-Memory Oracle storage systems are used for creating a *Smart Production*.
The monitored sub-division consists of eight separate, highly complex machines producing up to 1200 kg/h each, located in one interlaced division. About 2500 sensor data points are acquired in real time from the production process. The datasets are generated by about 30–500 sensor locations from the starting point to the end of each machine. A typical extrusion machine from Subsection 3.1.2 produces material within 20–120 min, generating 160 datasets each minute, as shown in Fig. 24. Each of the four to five machine PLCs (S7) is permanently connected via LWL, RJ45 to a SoftSPS data collector, linked to the machine network VLAN via ethernet (TCP/IP), communicating in milliseconds with one centralized Oracle database for data storage. The machines work in a *CPPS* communicating with several machine interfaces and data nodes. Up to 2500 process quality datasets are stored per minute, without considering growing data volumes or improving data quality. Each machine and sub-division is permanently linked to its production division, receiving additional data downstream from the Quality Management source. Attributes such as energy peaks, network routing performances, and file/server accession behavior are monitored continuously.

6.2.2 The Current Monitoring & Recommendation Layer

Each year, about 2 million separately SQL-queried datasets from the previous part of the Digital Factory are uploaded to the statistical software tools from the Oracle database for manual stacking by key attributes (e.g., time, order, unit, plant), for further data cleaning, which reduces the data amount to 21,900 datasets.
With certain targets in mind, the **Current Monitoring & Recommendation Layer** offers several different *Big Data* methods suitable for the exploration of the selected datasets and attribute prioritization by feature selection, for pattern recognition, cause sourcing, product identification, traceability, and simulation. Extensive statistical models are trained, evaluated and tested with datasets from one object, e.g., the machine comprising 20–90 datasets from 20–90 min from the years 2013, 2014, and 2015. The *State-of-the-art* methods, such as support vector machines, k-nearest-neighbor classifier, neural networks auto-encoder, and NOVCLASS, are modified to be used in the statistical models with underdetermined observation classes of unknown process conditions and sensor failures, and in the prediction of energy consumption behavior, explained in Chapter 4 and investigated in Chapter 7.
The examined results are reused in a second evaluation with a wider method spectrum (neural networks, support vector machines, k-nearest-neighbor classifier, naive Bayes, automated network search, and boosted trees) presented in Chapter 5. The software tools from Chapter 2, such as Matlab & Simulink, QuickCog, and STATISTICA, are used. The chronological prediction point is investigated between 1 and 2 min of analysis time, about 30–60 min in advance of the occurring condition problems for the selected machines.
The explored results from the interactive analysis and knowledge acquisition in the system architecture are implemented into an evaluated model within the same research environment, by the previously mentioned monitoring system, capturing performance measures (e.g., deviations) in real time. A traffic light system provides the status of the object or group of objects and their predicted behavior in advance, realized in two steps: (1) **Visualized Violation** and (2) **Recommendation System**.
The recommendation part of the standalone application for real-time condition monitoring analyzes condition changes and trajectory behavior, as explained in Chapters 4 and 5. Furthermore, it gives advice via interface by traffic light, sound, and by e-mail, to provide the responsible person with information on changes, counseling each machine operator to perform specified actions, as explained in Chapter 9. The Visualized Violation is traceable for up to 7 days. The standalone application is compiled in Matlab & Simulink[23]. The monitored attributes are reported automatically by business objects each day at 6 am for further processing. The machine operator is able to investigate the elapsed condition states to get an impression of the overall situation. Afterwards, the responsible process expert obtains the quality-related information by a generated pdf report. For further development, the interface will be extended to a multi-stage monitoring system, predicting condition changes from complete processing lines consisting of three machines in sequence, and for complete divisions of up to ten machines and for the plant.

[23] Statistical Toolbox and Compiler Toolbox necessary

6.2.3 The Current Adaptation Layer

The **Adaptation Layer** builds on the previous achievements, leading from visualized and conveyed recommendation, as presented in the previous subsection, to *Self-Configuration*. Therefore, the PLC data acquisition settings need to be changed from 'acquiring' to 'uploading' into the different PLC types. The on-line monitoring system obtains the right to take actions, e.g., in advance of occurring problems, by lowering the sensor location set value ("Extruder C – Temperature Zone 1" from 190 to 180°C) or by setting the actuators ("Torque 1" from status 2 to status 1). The best spectrum of settings needs to be stored for each product, considering one machine and the related waste reduction.
Regarding the *Self-Optimization* combined with energy efficiency targets, the best settings are taken for energy reduction, or for the *Self-Organization* terms with a focus on swarm approaches and multi-agent behavior. The *Self-Healing* part can be used to adjust to diverse influences, which has a high impact on quality entry topics regarding external material disturbances. The *Self-Protection* part becomes more important with regard to *IT-Infrastructure* safety issues, by focusing on security holes or breaches to detect foreign activities. Besides that, *Self-Explaining* is closely related to *Context Awareness* because objects with such properties are aware of their environment and of themselves. Such so-called intelligent objects learn capabilities from each other and understand how to interact with each other. Furthermore, the *Context Awareness* can be used for the direct interaction of objects with their environment (e.g., smart buildings recognizing human activities and adjusting intelligent light systems) [[F+13] [K+13]].
Full automation will be reached by autonomous guidance of the objects and their orientation towards the main business objectives of the division, plant, or corporate group. It is described as a self-managing system defined by some of the previously named properties (*Self-Configuration, Self-Optimization, Self-Healing, and Self-Protection*), which means that, e.g., the plant control station autonomously manages the monitored machines, their sensor locations, actuators and environmental influences, with additional equipment or maintenance tasks to prevent deviations in the productivity ratio [[KC03] [P+03]].

6.3 Current System Architecture

The system implementation overview with regard to the previously explained particular layers – the Acquisition Layer (A), the Monitoring & Recommendation Layer (B), and the Adaptation Layer (C) – are currently applicable to the mentioned object (machine) shown in Fig. 51. The conceptual design with the current implementations marked in green is displayed on the left; on the right side, the detailed architecture with the also green-marked currently implemented layers and the grey-shaded future approaches is presented.

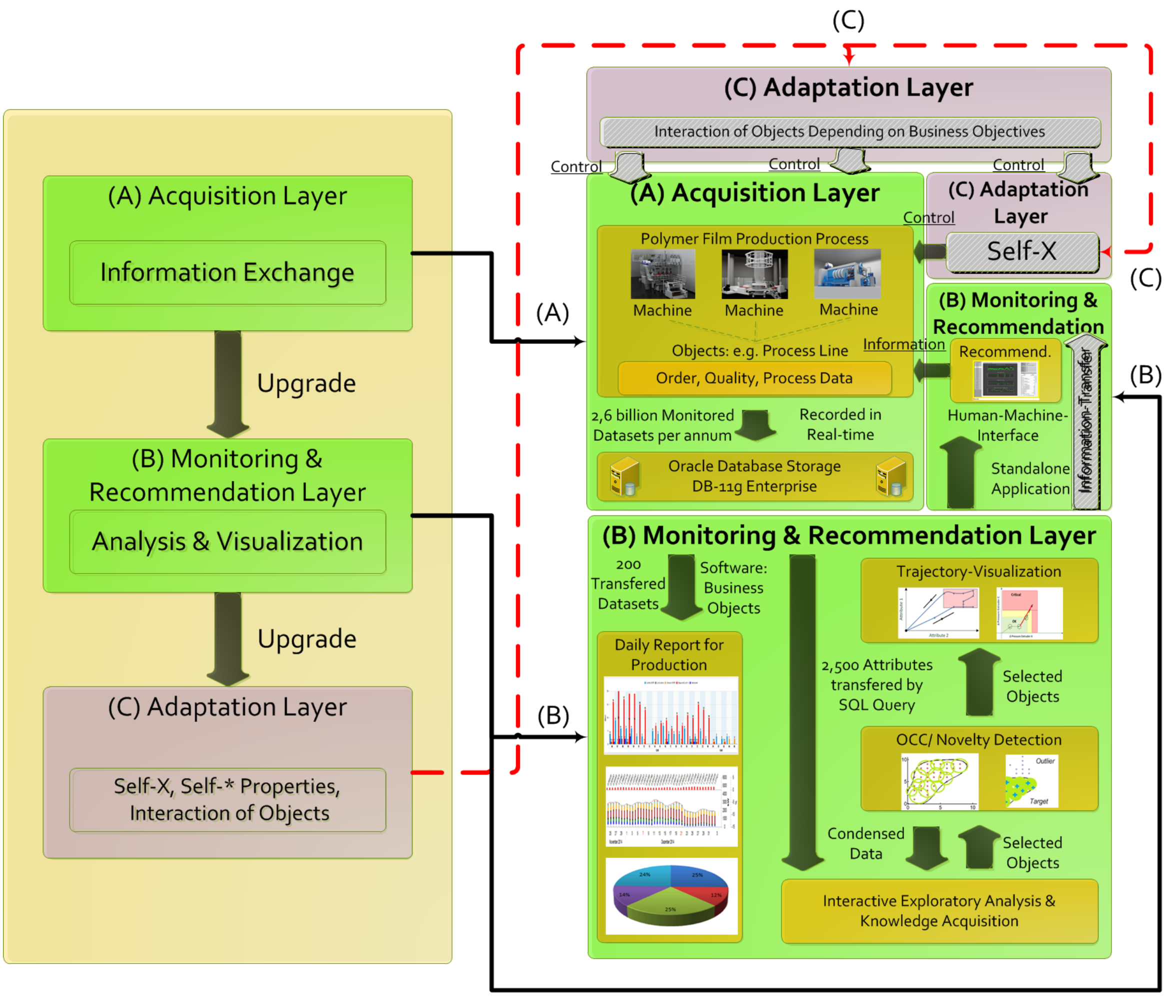


Figure 51 System architecture for the real-time computer-aided manufacturing processing action & recommendation system, consisting of implemented layers, marked green, and further investigation parts, shaded grey

In this case, first regarding the **Acquisition Layer (A)**, a *Smart Production and Digital Factory* has to be established. The acquired 21,900 datasets from multiple PLCs of, e.g., eight extrusion machines are then transferred to the Oracle database through the TCP/IP local area network and queried by the software tools, STATISTICA, Matlab, and QuickCog, for further investigation. This part describes the network and processing structures from Sections 2 and 3 that surround the machinery of the plant. The realized software workspace settings offer update functions for cyclic reloading of new datasets. Such direct connections form the block gateway from **(A)** to the **Monitoring & Recommendation Layer (B)**, with transferred information.

The feature selection by experts and QuickCog prepares the huge amount of datasets for the next examination. A downstream allocation separates the Big Data types into groups, to be explored in Subsection 7.1.1. Each tool provides an extensive variety of *State-of-the-art* methods (see Section 4.2: neural networks, support vector machines, k-nearest-neighbor classifiers, decision tress, naive Bayes, or modifiable methods; see Section 4.4: *One-Class* support vector machines, *One-Class* k-nearest-neighbor classifiers, *One-Class* neural networks, NOVCLASS) for *Big Data* analysis, predictive condition recognition (beside basic statistics), and extended optimization by *Trajectory-Visualization* approaches (described in Chapter 5) with the previously named software

tools.
In the next steps, the methods have to be investigated, which is presented in the following Chapter 7, and transferred to a *Human-Machine-Interface*, where the evaluated results are communicated between the machine and the human resources to provide an added value for the production department. The dynamic exchange can be achieved by various instruments, by messages, sounds, workflow changes, or typical visualization techniques, embedded into an open-loop monitoring and recommendation system (presented in Chapters 8 and 9). With regard to the *Knowledge-Discovery-Process* from Chapter 2, the previously examined insights have to be generated by standalone application monitoring cycles in real time, which means no more than 1 min of analysis refreshing time per machine, reporting, monitoring, and assistance. For research developments, a laboratory off-line monitoring system that watches one exemplary production machine offers the best opportunities for testing in a closed environment while not influencing the production system.
The interface structure should be precisely defined by the machine operators and extrusion experts, in iterative steps. Afterwards, the research development can be transferred to an on-line machine interface for a complete sub-division of, e.g., eight machines.
Additionally, for assistance developments, a discussion on a multitude of requirements asserted by the machine operators and their supervisors is required at the beginning of the case study, focusing on (a) the message transfer system, (b) the recommendation messages, (c) the message repetition, and (d) the interface structure and visualization types.
(a) The message system needs to be improved from a simple user interface with a message box to e-mail and SMS transfer, (b) the recommendation grades should be categorized into 3–5 graded text and traffic light types, (c) the repeated alarm should be set from fixed (each minute) to adaptable for each user between 1 and 120 min, by text, traffic light, or silence, and (d) the interface should offer several fixed time series plots, selectable machines, and the most important current attribute values.
The complete system from data acquisition to the monitoring interface can be improved by the involvement of human resources, such as technical *Industrie 4.0* experts and machine operators, in several steps, regarding different feedback issues, discussions, and SharePoint exchange types provided by the operators and employees. Their machine adjustment advices are directed to the recommendation part, which afterwards is provided by the installed system. As soon as quality deviations start to occur and the incident is foreseeable, after getting a real-time message, the monitoring interface can be checked by the supervisors and a "shut-down" can be initiated, which stops the machine for cleaning. This way, material waste could be verifiably reduced and the machine operators could be faster briefed on processing problems. With regard to the previously mentioned *Self-Monitoring*, a monitoring analysis cycle for the tolerance range of the current attributes over time would be an additional benefit, such as a comparison of the behavior within the last year, the last month, or the last 7 days until now.
Regarding the proposed system architecture from Section 6.1, for the Real-Time Computer-Aided Manufacturing Processing Action & Recommendation System, the particular blocks **(A) Acquisition Layer and (B) Monitoring & Recommendation Layer** have been implemented. The grey-shaded proposed extensions for future approaches, i.e. the **Adaptation Layer (C)**, replace the recommendation system in future investigations subsequent to this work by *Self-X* using PLC control systems, which are able to modify machine settings autonomously via wireless direct PLC access,

or make a diagnosis of, e.g., actuators, coordinated by centralized control stations. They display globalized monitoring data on dashboards and, by themselves, adjust the productivity ratios for products, machines, divisions, and plants for the corporate group worldwide.
The interaction between each level (e.g., **(A)** to **(B)** level network connectivity) and the data quality has to be locally observable in a decentralized manner, by machine-integrated monitoring functions that record connection interruptions and send alarm messages, and centrally, by the Oracle database supervising the subordinated databases. Connection errors have a high impact on the real-time action behavior of such approaches and attain main importance for the **(C)** exchange, regarding *Self-Configuration* failures due to missing datasets. In the following Chapters 7–9, the focus is laid on the **(B)** layer, consisting of the following experimental parts (7–9): (Chapter 7) The previously mentioned methods have to be inspected with regard to sensitivity, robustness, and flexibility for industrial applications, in view of the challenged targets. After further multi-stage investigation of the best experimentally tested method settings for the specific datasets, the evaluated system will be prepared for implementation in a dynamic quality loop for early process prediction, as already introduced in Chapter 5 and experimentally explored in Chapter 7.
After completion the complete setup in Chapter 8 for first testing – consisting of data acquisition (A) by the system architecture, the dataset types, and the best resulting methods – has to be investigated for its capability of integration into a standalone application (B). Therefore, the above-mentioned monitoring interface settings, message/recommendation options, update functions and exchanged improvement proposals have to be evaluated for compilation to an off-line laboratory Monitoring & Recommendation prototype system for one particular selected, not directly productive, connected machine (described in Chapter 8). Demands from the users need to be examined, such as the purpose of the system, refreshing of the loop, and the recommendation types.
After completion, the setup in Chapter 9 – and its improvement to a fully integrated *IT-Infrastructure* application to achieve an added value – has to be reproduced for installation into the productive polymer industry process. Thus, the extended data acquisition (regarding more query routines), the dataset types (differing PLCs), and the used methods (higher data volumes and slower processing times) need to be examined and validated. Furthermore, the interface settings have to be analyzed with regard to complexity reduction; the communication part has to be investigated concerning message types, and the interactive visualization has to be revised after information exchange with the particular machine experts for final acceptance (see Chapter 9).

7 Method Assessment and Optimization

In this chapter, the method assessment and optimization corresponding to the current system architecture implementation layers detailed in Chapter 6 is investigated. According to Chapters 4 and 5, the classification methods from Section 4.2 were studied with regard to the challenges described in Section 3.3.
The *OCC Selection and Evaluation* detailed in Section 7.1 explains the data acquisition and processing steps for a selected *cast* extrusion machine as well as the used types of off-line datasets, their sensor locations, and the applied feature reduction methods. Knowledge acquisition by modified *State-of-the-art* classification methods from previous blocks for testing the process condition recognition of new objects is evaluated, e.g., *one-class* support vector machines, *one-class* k-nearest-neighbor classifier, *one-class* neural networks (auto-encoder), and NOVCLASS.
The subsequent trajectory extraction (see Section 7.2, also introduced in Chapter 5) is based on the examined OCC Selection and Evaluation results for the machine in a second step for further optimization. The *Trajectory Prediction and Optimization* process cycle is a second, downstream classification step for testing with the *State-of-the-art* methods described in Section 4.2; it is extended to the third research approach, i.e. the prediction and reduction of the plant energy consumption described in Section 7.3.
Additionally, the involved process experts and machine operators were interviewed and asked to express their ideas of a monitoring system for the visualization of final recognition results, an interactive action & recommendation support, and the machine field for testing and implementation (see Section 7.4).
A concluding discussion about all examined results and further investigations is accomplished in Section 7.5.

7.1 OCC Selection and Evaluation

Multi-class classification methods were modified and extended for anomaly recognition of real production datasets. For process state prediction and process yield optimization, different modified *One-Class* Classification methods, as *One-Class* support vector machines, *One-Class* k-nearest-neighbor classifier, *One-Class* neural networks (auto-encoder), and *NOVCLASS*, presented in Chapter 4, were tested off-line.

7.1.1 Data acquisition and Extraction

The data acquisition and processing part is divided into, *(1) Data Extraction, (2) Training, (3) Validation, (4) Testing, (5) OCC Prototype, and (6) Online Novelty Detection*, for the specific use case from the polymer film industry, i.e. the film extrusion process at machine M150.

Data Extraction

The regular production of one product for up to 200 min, called roll, depends on the speed, which is manually adjusted by the staff. About 160 sensor locations, distributed from the beginning to the end of the extrusion process, are monitored within each minute at one machine.
In Fig. 52, the sensor distribution for the small extrusion batch of the polymer cycle is presented from the process start, with the starting point at 0 min; the diverse sensors acquire data on, e.g., MES output, speed, status, or energy consumption, as a small

selection out of 40 attributes. The dosing attributes are collected from about 5 min after process start, with 15 attributes[24].

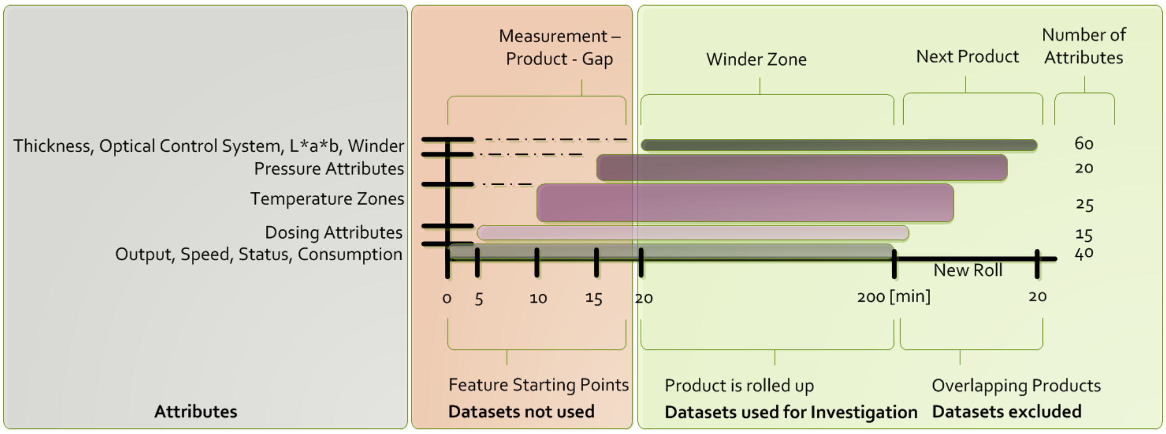


Figure 52 Temporal extrusion feature starting points for typical lower dimensional production process

In the temperature zones, the specified polymer data with 25 attributes are collected after 11 min of elapsed time, followed by 20 pressure attributes after 15 min and 60 winder and measurement attributes after 21 min. The transition time between the polymer material input and the final roll up is called "measurement–product gap"; it is caused by the processing length of about 15 m. The extracted part includes sparse datasets that belong to the current product whereas most of the data values are related to the previous product. Therefore, due to the overlapping sensor data, the subsequent 180 min are used in the below-described examinations, and not the first 20 min or the following 20 min relating to the next product.
In Table 9, the number of sensors at different machine parts in the production process is displayed in relation to their passing time. The extrusion product is rolled up in the winder zone until about 200 min have elapsed; it is then manually lifted and transported to the next interstation or warehouse, while the next roll production starts. The number of sensor points and their locations depend on the machine type and its

Table 9 Temporal matrix: time versus sensors with regard to the occuring number of sensor points

Sensors: / Time:	Output, Speed, Status, Consumption	Dosing Attributes	Temperature Zones	Pressure Attributes	Thickness, Optical Defects, L*a*b
0 – 4 min	# 40	–	–	–	–
5 – 10 min	# 40	# 15	–	–	–
11 – 14 min	# 40	# 15	# 25	–	–
15 – 20 min	# 40	# 15	# 25	# 20	–
21 – 200 min	**# 40**	**# 15**	**# 25**	**# 20**	**# 60**

[24]dimensions

products. As regards the product types, different aggregates are integrated into such machines, e.g., additional slitting equipment at the winder with 10 more attributes, if included.

The datasets are graded chronologically and, despite the gap, still offer valuable information, due to about 100 rolls produced in sequence. This means, however, that the first roll produced has a partial monitoring idle time of about 20 min. Quality changes occur with forewarning; thus, earlier recognition should be possible and aimed at. The amounts of bootstrap information rise during the first minutes, which is of high importance for further investigations.

The collected heterogeneous sensor datasets, e.g., pressure, speed, and the level of the water tank, from the programmable logic machine controllers are stored in an Oracle database, introduced in Chapter 3.

A particular selection of datasets comprising 21,900 min and 160 dimensions, each corresponding to a sensor point from the extrusion process including normal and conspicuous data, was taken from 2 years of run time (2013 and 2014) of the extrusion machine process. The screening was performed by the manufacturing execution system, by randomly selecting product rolls und thus covering production periods with normal and abnormal quality. Afterwards the processing datasets of these 250 product rolls were queried from the Oracle database by STATISTICA, Matlab and QuickCog for downstream off-line process analysis, as described in Chapter 2.

The production process comprises *State-of-the-art* quality information systems, such as in-line control systems (90% accuracy with an uncertainty of 0.2%, according to customer specifications, OCS GmbH [OCS15], and GUM [JCG08]), laboratory quality measurement systems (95% accuracy with an uncertainty of 0.1%, according to customer specifications and other laboratory equipment suppliers), and the operators' experience (unknown accuracy), for achieving objective conclusions and shall be extended by a standardized OCC monitoring system for unknown conditions; this cannot be achieved by the established information sources due to their location at the end of the process (Fig. 53). The current product roll quality monitoring by optical control systems and thickness measurements is located at the process end, i.e. at the winder, which is the last point of the extrusion production process.

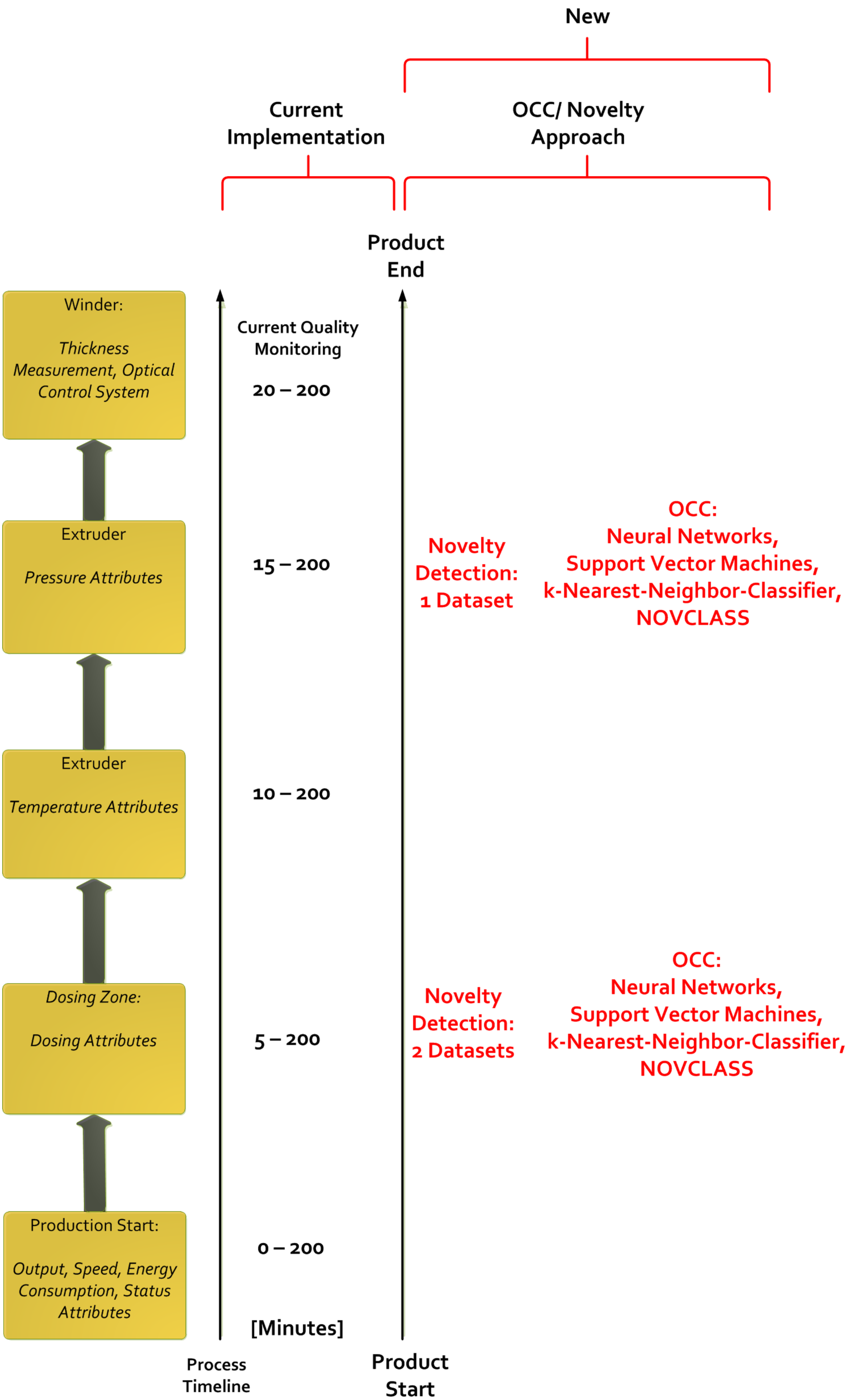


Figure 53 Current implementation and examination points from product start to end

The OCC approach aimed at finding an earlier quality prediction point within the product process, upstream of the currently implemented optical control system at the product end.
In a first-cut off-line analysis, experts identified those sensor locations (1–160) with high influence on the processing quality result, e.g., sensor no. 14 located at the "dosing part" in the extrusion line, anonymized in Table 10 due to confidentiality reasons.

Table 10 Sensory data selection by experts and QuickCog, basic method for feature reduction

Experts Sensory Data Selection	QuickCog Sensory Data Selection
11, 18, 21, 23, 24, 25, 28, 54, 55, 56, 70, 72, 74, 76	8, 24

The feature reduction with a basic method in QuickCog (sequential backward selection, sequential forward selection, qs separability) showed sensor points 8 and 24 (pressure_C1 in extruder[25] C and pressure_A1 in extruder A) as having the most significant impact on the process deviations, whereas the process experts did not assign any important influence to sensor point 8 because extruder C is a by-pass extruder to the main extruder A. However, the pressure_B1 attribute, which was similar to pressure_A1, and three out of these calculated attributes (pressure-calc_A2, pressure-calc_B2, pressure-calc_C2) were additionally selected as highly important for the following supervised approach on a multi-class based selection.
Regarding the downstream quality rate, the main attributes for process state description then resulted in 6 pressure attributes (*pressure_A1, pressure_B1, pressure_C1, pressure-calc_A2, pressure-calc_B2, pressure-calc_C2*) regarded in Dataset 1. Vulnerability to erroneous sensor properties of the dosing part within the extrusion line is described by two exemplary datasets consisting of four dosing attributes each (*Dosing 1, Dosing 2, Dosing 3, Dosing 4*, and Dosing 5, Dosing 6, Dosing 7, Dosing 8) from two machine parts. The applied features could lead to classification errors in the future, due to a missing repeatedly implemented selection cycle. The feature selection could have been extended by other methods as, e.g., correlation of normal and abnormal conditions, or feature screening with additional procesing software types [[KK12] [M+01] [Kö01]].
The three extracted experimental datasets are specified in Table 11. Dataset 1 describes the process conditions (such as pressure values) and their faulty states; Datasets 2 and 3 refer to missing or erroneous dosing data due to sensor problems.

[25] Aggregate for melting up PE

Table 11 Extrusion line off-line dataset selection for OCC methods and their description

Dataset	Type	Count	Dim	Description
Dataset1 (Normal)	Rigid Film Pressure Raw Data 1 (EPD1_Feature) – Pressure Data – Normal	13215	6 of 160	Datasets without faulty data (normal data)
Additional to dataset 1 (Abnormal)	Rigid Film Pressure Raw Data 1 (EPD1_Feature) – Pressure Data – Abnormal	2055	6 of 160	Datasets with faulty data (abnormal data)
Dataset2 (Normal)	Rigid Film Dosing Raw Data 1 (RFDD1_Feature) – Dosing/Mixture Data – Normal	2114	4 of 160	Datasets without faulty data (normal data)
Additional to dataset 2 (Abnormal)	Rigid Film Dosing Raw Data 1 (RFDD1_Feature) – Dosing/Mixture Data – Abnormal	765	4 of 160	Datasets with faulty data (abnormal data)
Dataset3 (Normal)	Extrusion Dosing Raw Data 2 (RFDD2_Feature) – Dosing/Mixture Data – Normal	953	4 of 160	Datasets without faulty data (normal data)
Additional to dataset 3 (Abnormal)	Extrusion Dosing Raw Data 2 (RFDD2_Feature) – Dosing/Mixture Data – Abnormal	667	4 of 160	Datasets with faulty data (abnormal data)

Each dataset type is represented by normal, datasets without faulty data, and sparsely countersampled abnormal data, datasets with faulty data, specifying different conspicuous process problems.

7.1.2 Data Processing

To ensure that the training and test data are properly separated for further investigations, the datasets are split into training and testing sets, and the testing data, additionally into validation and testing cases. The exact partitioning of the datasets was experimentally tested with the k-fold cross-validation (beside the leave-one-out) method, on a small random selection of 500 observations from Dataset 1 to avoid overfitting.
The method was tested with the k-nearest-neighbor classifier to examine the quality of the separation. The classification trees are also suitable for such approaches. The

separation quality was measured by the resubstitution loss[26] and the cross-validation loss[27] when predicting based on data not used for training. The k-fold cross-validation for k = 2, known as hold-out method, separated the datasets into k subsets with different sizes, and selected one for testing and the rest for training. No further result improvements by different k-fold settings, such as k = 10, were achieved. For the spot test, the exemplarily chosen k-nearest-neighbor classifier method with $k = 2$ was used. The average misclassification loss (error) was calculated for the subset. Fig. 54 displays the resubstitution loss, the fraction of misclassifications from predictions for training datasets, the resubstitution time for calculation, and the average cross-validation loss defining the loss of each cross-validation model on datasets not used for training.

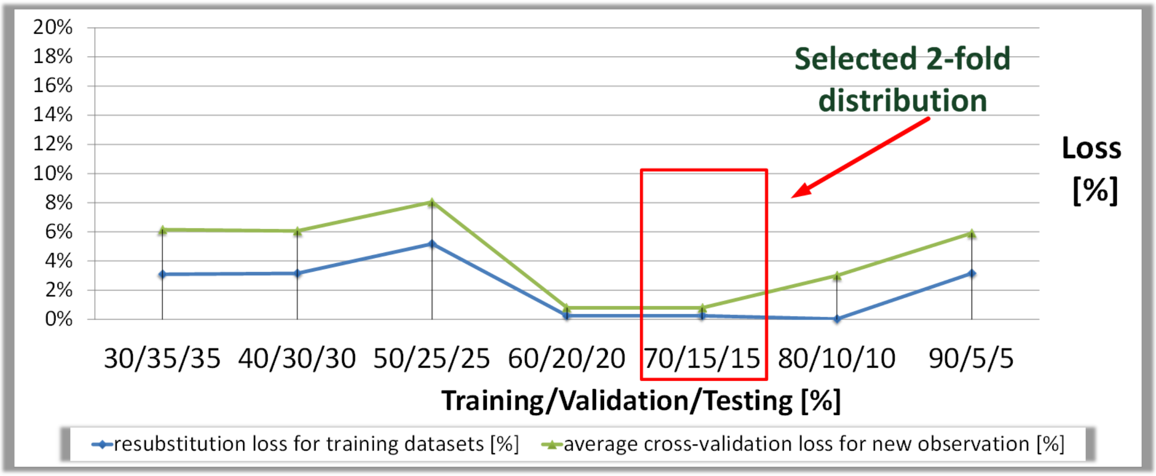


Figure 54 2-fold cross validation results for dataset separation into training, validation, and testing data (hold-out method)

The spot-tested hold-out method with data separations into training, validation, and testing sets of 60%, 20%, 20% and 70%, 15%, 15% showed the best average cross-validation loss results for new observations of less than 2%. In the following, a dataset separation of 70%, 15%, 15% (training, validation, testing) was chosen to provide a higher amount of training data.
The next step divided each dataset (Datasets 1, 2, and 3) into the previously examined distributed cases for downstream model building with the modified OCC classifier types described in Chapter 4. In the following, the performance measurement, the sensitivity analysis for the general OCC methods, and their recognition accuracy are explained.

7.1.3 Results Examination

The previously defined, acquired and partitioned three datasets from upstream locations, the dosing zone (RFDD1_Feature, RFDD2_Feature) and the extruder (EPD1_Feature), were investigated with four OCC methods: the *Neural Networks (Auto-Encoder), Support Vector Machines (RBF), the k-Nearest-Neighbor Classifier, and NOVCLASS*. For the performance measurements of the methods, the mean specificity and the standard deviation were calculated.
The predicted performance measurement, as described by ROC curves in Chapter 2 (rate of false positives (FP) against rate of false negatives (FN)) can be divided into four

[26] Fraction of misclassification
[27] Average loss of each cross-validation model

categories including true positives (TP) and true negatives (TN), as shown in Table 12.

Table 12 Classification performance measurement in truth table confusion matrix

	Positive	Negative
Predicted Positive	TP	FP
Predicted Negative	FN	TN

In the following the TP, and TN are regarded, depending on the respective normal or abnormal class, which is displayed as specifity:

$$'Specificity = (TN)/(FP + TN)' \tag{3}$$

Additionally, the precision (number of TP divided by the number of TP and FP), the accuracy (number of TP and TN divided by the number of P and N), the sensitivity (number of TP divided by the number of TP and FN), or the F1 score (harmonic mean of precision and sensitivity) can be applied, depending on the count of necessary ratio types for prediction.

The modification of different preselected methods for optimized novelty detection was investigated on predefined datasets, consisting of data collected over 2 years (2013 and 2014), defining sensor properties from the dosing stage (RFDD1_Feature, RFDD2_Feature) and operating condition machine properties (EPD_Feature) from the extrusion stage.

The settings for, e.g., the k-nearest-neighbor classifier (k = 1, Rej = 0,1; 0,01; ... / k = 2, Rej = 0,1; 0,01; ...) were consecutively modified and repeatedly minimized to the best results (after 10 runs), as shown in Table 13. For example, the best results from

Table 13 Sensitivity analysis for general OCC methods

Methods	Classifier Type	Best Settings
Neural Networks (Auto-Encoder)	OCC	HiddenNeurons = 8; Rej = 0.00001;
Support Vector Machines (RBF)	OCC	Sigma = 5 and 10; Gamma = 0.1; Rej = 0.001;
k-Nearest-Neighbor Classifier	OCC	k = 2; Rej = 0.001;
NOVCLASS	OCC	ScaleFactor = 1.3 ;

the NOVCLASS method are illustrated with a scale factor of 1.3 for the given datasets, and from the Neural Networks Auto-Encoder for Dataset 1 with the normal operating conditions and for the abnormal class with critical operating conditions.

As shown in Table 14, the best settings for the Auto-Encoder Neural Networks were achieved with 8 hidden neurons and a rejection rate of 0.00001, leading to the deterministic empirical, lowest repeatedly received specificity rate of 96.4% for Dataset 1 after the execution of 10 runs, with an experimental standard deviation (STD) rate smaller 2.5%. The experimental datasets for the normal conditions from Fig. 11 were executed by hold-out for off-line training (70% of the data), validation (15%), and testing (15%). The normal conditions include the following process behavior: $<$ 5% waste, output of $>$ 1000 kg/h, fast production with $>$ 100 m/min, energy efficiency of $<$ 0.6 kWh/kg, low maintenance cost, best roll quality with 0 incidents, and no additional human resources

Table 14 Experiments' specificity with 4 OCC tested methods, and 3 datasets (only normal datasets used)

	Methods OCC	Dataset	Condition Class	Mean in [%]	STD in [%]	AbsErr [#]
1	Neural Networks (Auto-Encoder)	Dataset 1	Normal	97.6	0.4	317
		Dataset 2	Normal	99.1	0.1	19
		Dataset 3	Normal	99.5	0.4	5
2	Support Vector Machines (RBF)	Dataset 1	Normal	99.9	0.1	13
		Dataset 2	Normal	99.6	0.2	8
		Dataset 3	Normal	99.8	0.1	2
3	k-Nearest-Neighbor Classifier	Dataset 1	Normal	98.8	0.2	159
		Dataset 2	Normal	98.7	1.1	27
		Dataset 3	Normal	99.9	0.1	1
4	NOVCLASS	Dataset 1	Normal	99.8	0.1	26
		Dataset 2	Normal	99.8	0.2	4
		Dataset 3	Normal	99.4	0.1	6

necessary in contrast to abnormal states. In comparison, the abnormal conditions lead to the worse results, with low productivity due to high rework and capacity problems for the divisions.
The requirement in the polymer film industry for efficient control through a recognition system is 90% for the specificity rate, which defines the error of the target class to become an outlier; this specificity is similar to that of the optical control systems for defect detection, set up in cooperation with clients. All OCC methods achieved more than 90% specificity. The elapsed time for recognition is 13 s for one machine, and about 69 s for seven more machines, depending on the number of attributes to be analyzed. In the following, the next step, from training, validation, and testing to an OCC prototype, is examined.

7.1.4 OCC Prototype and Online Novelty Detection

Based on the results from the dataset separation and investigation by OCC methods previously examined off-line, a prototype visualization approach and on-line testing with an on-line data connection was investigated in the next step. To this end, in a first approach, Dataset 1, i.e. the extrusion pressure dataset 1 (EPD1_Feature), was selected for on-line visualization, due to the direct location at the main extrusion machine processing part: the extruder.
To reproduce the 6-dimensional dataset in a 2-dimensional space for visualization, dimensionality reduction methods were applied, such as multidimensional scaling [BG05]. Approaches for dimensionality reduction are divided into linear (principal components analysis and classical multidimensional scaling (MDS)) and non-linear types (manifold learning, SOM, and non-classical MDS). The multidimensional scaling approach – beside self-organizing maps, parallel coordinates, and principal components – was chosen due to similarities of the previously used feature reduction methods with QuickCog.
In the following, two different types of multidimensional scaling were distinguished: the classical (linear, metric) [[Tor52][Tor58]] and the non-classical (non-linear, non-metric

= ordinal) scaling type [KW78].

Distance Type

First of all, before the dimensionality reduction part was started, a distance matrix had to be created from the pressure observations. For this purpose, the different types of distances were tested to form a distance matrix including all unique observation points from the pressure dataset in a first step. Their computation, according to [[The15] [F$^+$96]] is shown in Table 15. For instance, the Euclidean type computed a distance

Table 15 Distance types and computation, according to [The15]

Type	Computation
Euclidean	$d_{kt} = (\mathbf{x}_k - \mathbf{x}_t)(\mathbf{x}_k - \mathbf{x}_t)^T$
SEuclidean[28]	$d_{kt}^2 = (\mathbf{x}_k - \mathbf{x}_t)\mathbf{V}^{-1}(\mathbf{x}_k - \mathbf{x}_t)^T$
Cityblock	$d_{kt} = \sum_{i=1}^{n} \lvert x_{ki} - x_{ti}\rvert$
Minkowski	$d_{kt} = \sqrt[p]{\sum_{i=1}^{n} \lvert x_{ki} - x_{ti}\rvert^p}$
Chebychev	$d_{kt} = max_i\{\lvert x_{ki} - x_{ti}\rvert\}$
Mahalanobis[29]	$d_{kt}^2 = (\mathbf{x}_k - \mathbf{x}_t)\mathbf{C}^{-1}(\mathbf{x}_k - \mathbf{x}_t)^T$
Hamming	$d_h = \sum_{i=1}^{n} (x_{0i} \oplus x_{1i})$
Jaccard	$d_{kt} = \frac{\#[(x_{ki} \neq x_{ti}) \cap ((x_{ki} \neq 0) \cup (x_{ti} \neq 0))]}{\#[(x_{ki} \neq 0) \cup (x_{ti} \neq 0)]}$

between all pairs of objects from the m-by-n data matrix (columns m = attributes, rows n = observations), displayed in a vector arranged in the order (2,1), ..., (k,1), (3,2), ..., (k,k – 1). Afterwards the generated vector of length k(k – 1)/2 was converted into a squared n-by-n matrix form (distances between all n observations). Each element (ii,jj) in the matrix, ii < jj, corresponds to the distance between the objects ii and jj in the dataset.

[28] Each difference scaled by the standard deviation (std) - **V** is diagonal matrix with vectors of std

[29] **C** is the covariance matrix

Scaling Type
Now the next step, the scaling for dimensionality reduction, was possible. The **classical** scaling took the previously calculated squared n-by-n distance matrix **D** to generate a new configuration matrix **Y** and the eigenvalues Y*Y'. The interpoint distances were approximated to the distance matrix; thereby, the dimensionality reduction is achieved for visualization, e.g., p = 2 or 3 dimensions, according to the interval MDS, [BG05].

$$dd_{ij}(\mathbf{X}) = aa + bb * pp_{ij} \tag{4}$$

Other used cases are ordinal and monotone functions, which were not computed in this case. The error function is calculated as:

$$err_{ij} = f(pp_{ij}) - dd_{ij}(\mathbf{X}) \tag{5}$$

The approximation ratio for reduction to the wanted 2 or 3 dimensions, tested with the above-displayed distance types, was measured by a reconstruction error for multidimensional scaling, calculated for the chosen number of dimensions with different distance types, as shown in Table 16. The table offers an overview of the different distance

Table 16 Distance types reconstruction error (stress) for dimensions 2 and 3

Type	Dim 1	Dim 2	Dim 3	max(Distance)
Euclidean	0.21	0.08	0.02	40.00
SEuclidean	0.24	0.13	0.07	14.43
Cityblock	0.22	0.11	0.07	68.16
Minkowski	0.21	0.08	0.02	40.01
Chebychev	0.24	0.16	0.15	37.00
Mahalanobis	0.25	0.22	0.17	14.81
Hamming	0.37	0.34	0.33	4.00
Jaccard	0.36	0.35	0.34	4.00

measurement types, their reconstruction error transforming the 6-dimensional space into less, i.e. 2 and 3, dimensions, and the maximum distance between observations. A maximum relative error ratio close to 0 allows the best reconstruction results; higher ratios lead to poor results [$\mathrm{F}^{+}96$].
In this case, for the 2-dimensional view, the Euclidean Distance type showed a maximum relative reconstruction error of 0.08 whereas the maximum distance max(**D**) is 40. Therefore, the reconstruction is poor and, thus, in the 3-dimensional view, the reconstruction is good with 0.02.
The above-described measuring sequence in short: (1) calculation of the distance matrix (best choice Euclidean), (2) performing the classical multidimensional scaling, (3) measuring the reconstruction error (lowest reconstruction error for Dim 2 was Euclidean).
The **non-classical** multidimensional scaling used the squared n-by-n Euclidean distance matrix D for calculating an approximated transformation, depending on different

criteria, e.g., the non-metric scaling: the sum of squares and 4th powers for interpoint distances, called stress1 and squared stress1. Another criterion used was the metric scaling for dissimilarities in **D** as, e.g., the sum of squares and 4th powers of the dissimilarities, called stress and squared stress, furthermore the Sammon's nonlinear Mapping for positive off-diagonal dissimilarities. For the calculation of the initial points, the random starting point, the number of repetitions (200) of the initial configuration, and the termination tolerance for the stress criterion (1e-4) were chosen by spot test.
The interpretation of the criteria and their goodness of fit was provided by Kruskal (1964) [[Kru64a][Kru64b]], as shown in Table 17, used for a comparing investigation of the different scaling methods in the following.

Table 17 According to Kruskal (1964) the interpretation of the goodness-of-fit criterion

Stress	Goodness-of-Fit
0.2	poor
0.05	good
0.00	perfect

For each criterion, Kruskal (1964) defined an interpretation scale for decision help. The best performance was achievable with an assessment error of 0.00; good performance was obtained with 0.05, and poor reconstruction with 0.2 and higher.
The reconstruction error for the different scaling type criteria, such as stress, sstress, metricstress, metricsstress, and Sammon's stress explained above, is displayed in Table 18. The best results were achieved with the Sammon's Mapping stress [[H$^+$09][Sam69]]

Table 18 Non-classical multi-dimensional scaling goodness-of-fit (GoF) criterion 2-D

Criterion	Scaling Type	Iterations	2 Dim	GoF	1 Dim	GoF
Stress	non-metric	48	0.031	good	0.212	poor
SStress	non-metric	200	0.022	good	0.226	poor
Metricstress	metric	36	0.039	good	0.259	poor
Metricsstress	metric	41	0.028	good	0.272	poor
Sammon's Stress	metric	32	0.002	perfect	0.096	good

after 32 iterations and with the squared stress after 200 iterations, with perfect and good goodness-of-fit of 0.002 and 0.022, respectively. Results below 0.6 are generally found to be acceptable for a good fit between the current dataset and the number of regarded dimensions. On average, the metric scaling achieved better results compared to the non-metric scaling criteria. Therefore, the non-classical scaling with Sammon's metric was chosen as best criterion in the following for the finally scaled 2-dimensional view (Fig. 55).

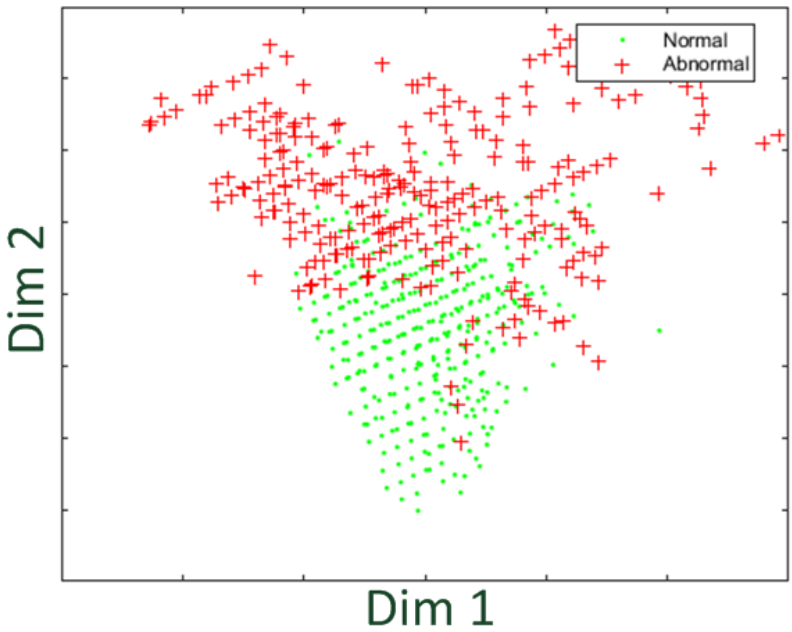


Figure 55 Sammon's projection for the stress of 0.002

The above Sammon's mapped view gives a vision of the scaled dimensions after 32 iterations. The distribution of the visualized dataset points comprises about 23% of the dataset points marked as a separate group of red crosses representing the abnormal conditions, and about 5% mixed red and green dataset points (most of them located very close to each other, which could be optimized by the weight-watcher function in QuickCog for uncovering mapping errors) representing the boundary part, and 72% of green dataset points representing the normal conditions. The numerous points occlude each other in the 2-dimensional visualization. The denotation of the scaled 2-dimensional view for one plot was changed later from "Dim 1" and "Dim 2" to "Pressure 1" and "Pressure 2", according to the demand of the staff for a better understanding of which data is monitored.
Different methods were tested to find the best fitting type according to the dimensionality reduction. The final criterion was used in the next steps to achieve the best possible visualized 2-dimensional view for further investigation.
The above-described measuring sequence is given in short by: (1) calculation of the distance matrix (best choice achieved by the Euclidean Distance), (2) performing the non-classical MDS, (3) measuring the goodness-of-fit (best choice achieved by the Sammon's metric).
In the following, the Sammon's metric non-classical MDS was implemented for further investigation.

Visualization

The applied Sammon's Mapping coded by Cawley (2007) [The15] offered the visualization of the final iteration (200). The multidimensional unscaled dataset projections were transformed into a 2-dimensional scaled view by Sammon's metric for staff interface visualization [[RD97] [Kö00]]. Afterwards, the display was separated into zones by Delaunay convex hull and threshold settings.
The transition from the first to the fourth sub-plot is described in more detail. In this

part, a Delaunay triangulation[30] [[Del34][The15]] was implemented to define all points on the outer line of the displayed data cloud (boundary), which is called the convex encircling hull.

In fact, the heuristic approach of the Delaunay triangulation directly applied to the 2-dimensional data to achieve the convex hull generates an imprecise grid, which still fits the requirements for appropriate representation of the region of interest but should be kept in mind as an error source. In the next step, the threshold was set by the maximum of Dimensions 1 and 2 of the normal dataset on the encircling hull (Fig. 56).

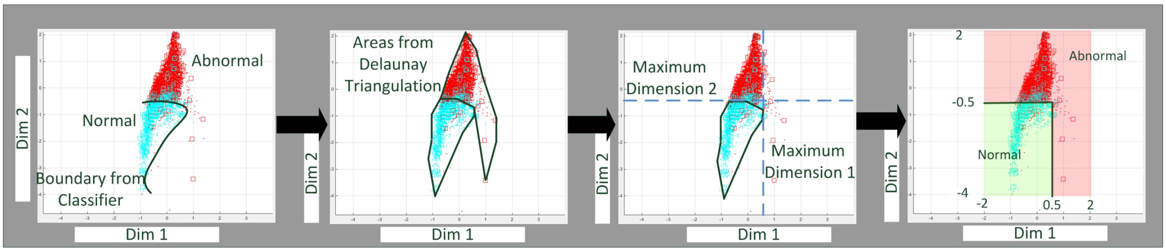


Figure 56 (1) Sammon's projection, (2) Delaunay triangulation, (3) threshold setting, (4) zone visualization

The following zone visualization was achieved in a heuristic approach as a kind of linear separation. On the one hand, every data point above the maximum of Dimensions 1 and 2 was visualized as abnormal part and, on the other hand, everything below the maximum, as normal condition.

From the final scaled view with two zones (normal, abnormal), an additional third zone was generated manually (Fig. 57). The normal zone *green*, the boundary zone *yellow*, which is part of the normal zone, and the abnormal zone *red*, described by hypersphere forms, are visualized in 2 dimensions but with a vector consisting of 6 attributes behind each data point.

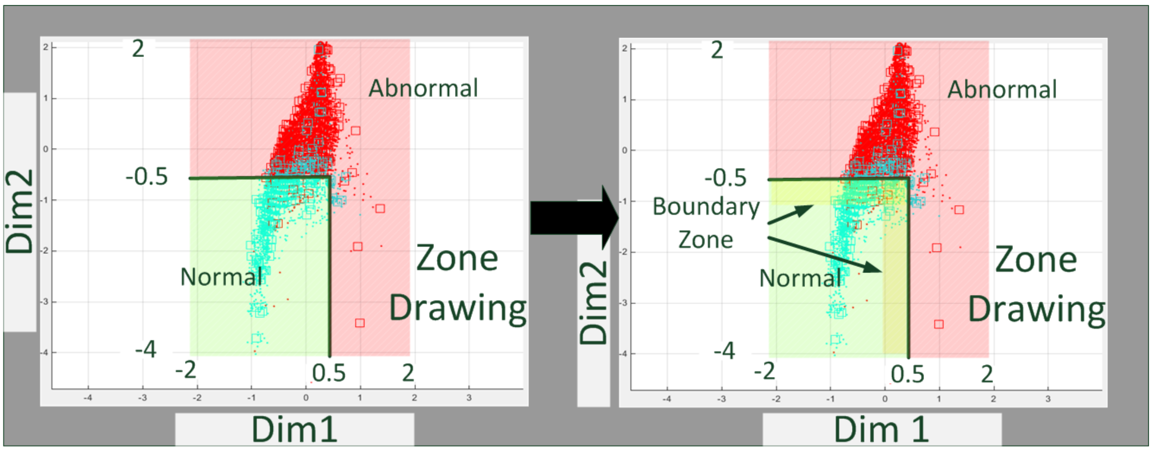


Figure 57 Zone drawing for normal and abnormal behaviour, and the integrated boundary zone for staff visualization

The boundary zone is part of the normal zone, but defines the region of conspicuous condition behavior, with a spot-tested risk of 20% for moving into the abnormal zone.

[30]Provided for convex hulls to make direct connections between the maximum and minimum points of a dataset

The boundary zone was additionally separated within the normal zone, integrated as a spot-tested 30% distance from the boundary. In this area, the process behavior could slowly change to slight quality incidents occurring for the first time, such as defects like *holes and gels*, which could lead to waste.
The complete cycle of the previously explained process steps from multidimensional scaling to the final visualized zones is displayed in Fig. 58.

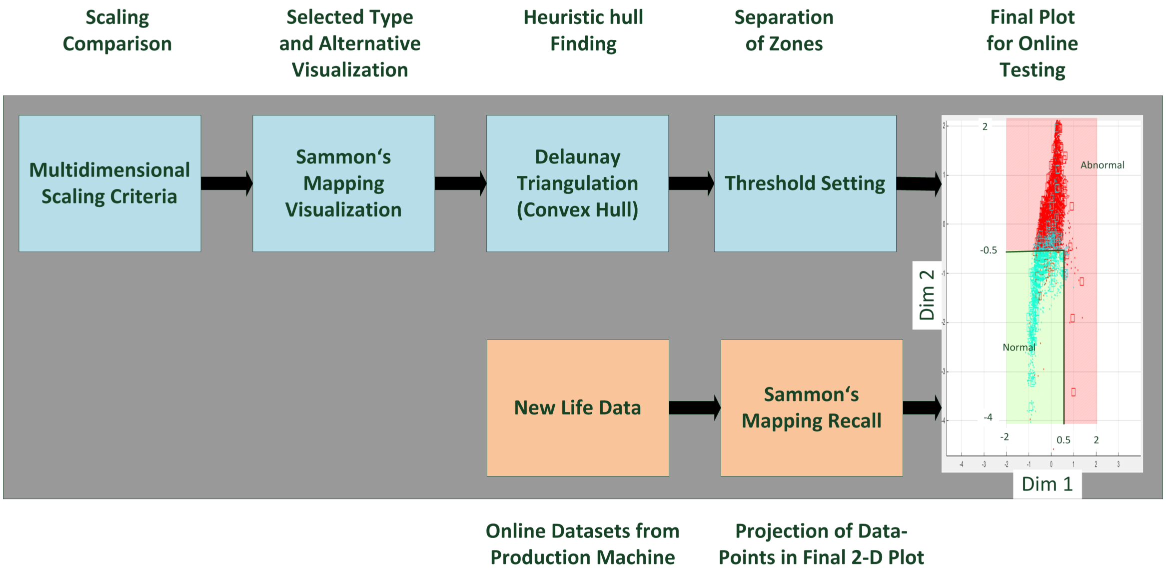


Figure 58 First process cycle: (1) multidimensional scaling (Sammon's projection), (2) visualization, (3) Delaunay/convex hull, (4) threshold setting, (5) final plot and second process cycle: (1) life data acquisition, (2) Sammon's Recall Mapping for projection into 2-D

The derived from the previous hypersphere *OCC Prototype* 2-dimensional final plot was used for downstream on-line visualization, according to staff interviews, explained in Section 7.4. According to the subjective demand of the staff for their better understanding of the meaning of the data monitored, the following demonstrative traffic light zone descriptions were set:

- Green: Normal Zone
- Yellow: Boundary Zone
- Red: Abnormal Zone

After off-line testing with samples of the 21,900 abnormal and normal (target) datasets, the scheme for Dataset 1 with regard to the results examination was plotted with the examined boundary from the results examination. For the two new scaled attributes, Dimensions 1 and 2, the 19,500 new datasets were extracted from the 2-dimensional reconstruction.
To this end, a Sammon's Mapping recall (non-linear mapping recall = NLMR) according to [Kö00] was executed to relocate the new datasets in the previously defined mapped 2-dimensional view. For the recall, the distances between the training patterns and the final mapped pattern were calculated and the transformation was applied to new online-acquired life datasets from the polymer production process. The new datasets were separately reconstructed to fit into the final mapped plot.
The procedure generates a stable training pattern, into which the new life datasets

are mapped with less time spent on repetition, which is an improvement compared to the initial instable mapping of all shared training and testing datasets. Only the new data points are processed this way and applied in the following procedure for online acquisition.
Furthermore, triangulation or similarity methods offer additional mapping support, which should be investigated in future studies. A parallel approach with an ensemble of neural networks suggested by Gianniotis (2013) [The15] and shown in the appendix was tested but not further investigated.

7.1.5 Complete Assessment Loop

The complete *OCC Selection, Evaluation, and Visualization* cycle starts with the data acquisition and extraction of 3 new datasets, reduced to 8 dosing attributes and 6 pressure attributes (Section 7.1.1), followed by data processing for the off-line distributed training, validation, and testing (Section 7.1.2) and results examination of the tested samples (Section 7.1.3), and ends with the OCC prototype and online novelty detection (Section 7.1.4) (Fig. 59).

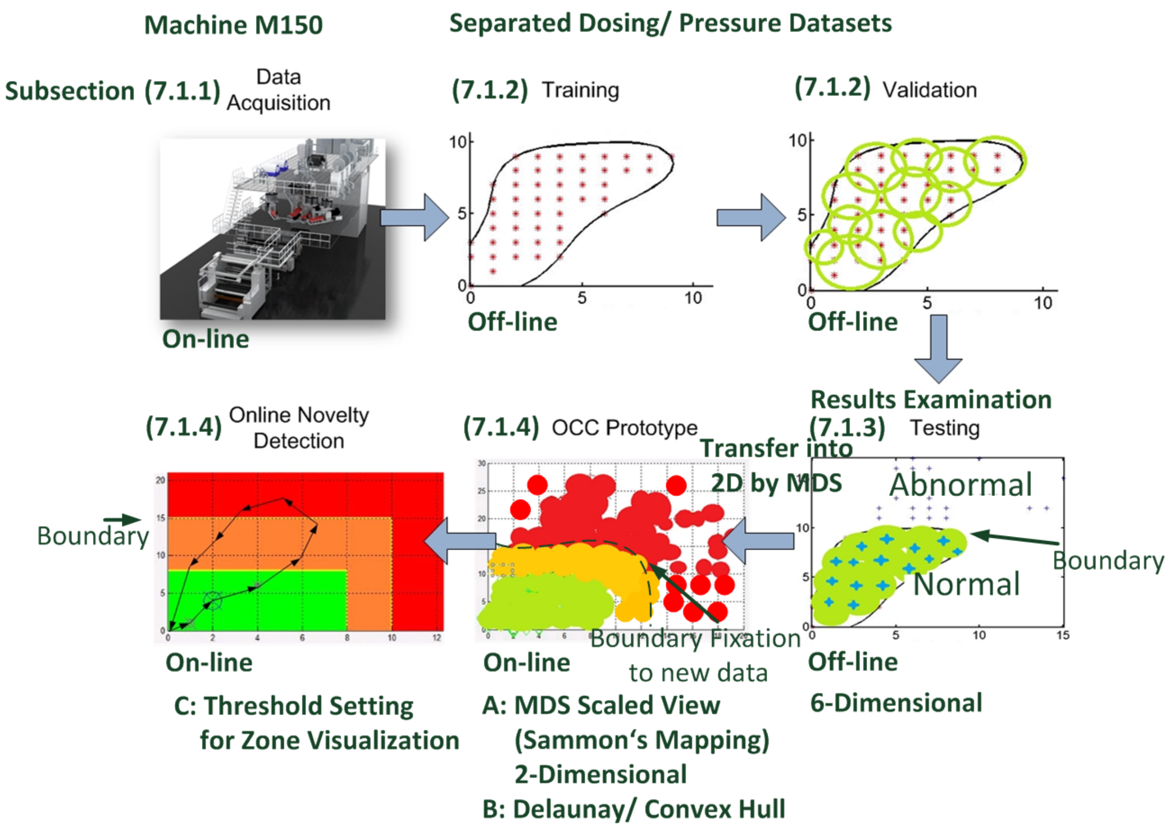


Figure 59 Process diagram for data acquisition, off-line training, validation, and testing, developed OCC preliminary on-line prototype for normal and abnormal datasets, and simplified on-line operators interface with sample process trajectory

The process diagram shows the data processing cycle for the monitored datasets of the examined product rolls. The transfer from the 6-dimensional construction into a lower, 2-dimensional view was established by a multidimensional scaled, Sammon's mapped

view, extended by a Delaunay convex hull heuristic adaptation and threshold settings, according to Fig. 56, for zone visualization in the final plot of the pressure dataset in focus. According to Fig. 57, the final plot was generated by threshold setting on the convex hull of the normal zone for separation of the zones.
In the following, the Pseudo-Code 10, for *OCC Selection and Evaluation* summarizes the previously described consecutive steps one by one:

begin

1. **Acquisition** of Dosing and Pressure Datasets (3th) from Oracle DB
   SQL Query to Database
   Collecting Normal/ Abnormal Datasets (fetch [...])
2. **1st Feature Reduction (160Dim → 6Dim)** with QuickCog
   Sequential Forward Selection
   Features (2) → sequentialfs(fun,observations,attributes)
   6Dim ← (1) Process Experts + (3) Calc. Features + (2) Features
3. **Separation** of Datasets into Training, Validation, Testing
   Cross Validation Applied
   Crossvalind('k-fold', N, 2)
4. **Model building with OCC Methods** Off-line 6Dim
   NOVCLASS
   Number of N samples and n features defined
   Winning neuron w_{ji}^{NN} of all N good sample vectors $\mathbf{w}_{kji}$
   $w_{ji}^{NN} = \min_{k=1}^{N} (\sum_{f=1}^{n} (\mathbf{p}_{ijf} - \mathbf{w}_{kijf})^2)$ [KG05]
   Determine the novelty distance $\mathbf{D}_{nov}^{ij}$ between the nearest neighbor neuron w_{ij}^{NN} and current pixel vector $\mathbf{p}_{ij}$
   Transform the distance $\mathbf{D}_{nov}^{ij}$ into a gray-level value $\mathbf{G}_{nov}^{ij}$
   $\mathbf{G}_{nov}^{ij} = f(\mathbf{D}_{nov}^{ij} - \Theta_{ij})$
5. **2nd Feature Reduction (6Dim → 2Dim)**
   Multi-Dimensional Sammon's Scaling Applied
   Sammon's Mapping
   Initial randomly generated lower-dimensional pivot vectors $Y(m)$ according to vectors X from original high-dimensional space
   Iterative Update of $Y(m)$ with steepest descent for min(Error)
   Error $= \frac{1}{\sum_{i<j}^{N} dX_{ij}} \sum_{i<j}^{N} \frac{(dX_{ij} - dY_{ij})^2}{dX_{ij}}$ [Sam69]
   Evaluation of the Sammon's Mapping error: $E_{new} < E_{old}$
6. **Visualization of 2-D Mapping**
   Delaunay/ Convex Hull Calculation
   Each 2-D Data Point extended by z- Coordinate $= x^2 + y^2$
   Convex Hull around all 3-D Points
   Projection of Triangles into 2-D Map
7. **Boundaries of 2-D Mapping**
   Threshold Setting
   Maximum Data Points for Normal Zone (Max(Dim1, Dim2))
8. **Projection Testing Data into 2-D Mapping**
   Sammon's Mapping Recall [Kö00]
   Projection of new Data-Points into 2-D Mapping

end

Pseudo-Code 10: Procedure of the OCC selection, evaluation, scaling and projection approach

7.1.6 Summary

The goal in Chapter 7 was a practical approach to and comparison of the modified OCC methods applied to real datasets from manufacturing sensor locations. The danger emanating from unknown operating conditions and sensor failures leading to material waste is traced back to an insufficient understanding of the process. Randomly occurring states from transmission problems, missing data, and novel machine conditions (due to environmental influences) reflected in more than 100 process attributes exceed the operators' experience and the standardized process control. Anomaly-seeking methods for advanced examination are necessary to describe the sensor behavior and process properties in order to improve the recognition of abnormalities.
In this part, four viable OCC methods (e.g., *One-Class* support vector machines) were investigated with different datasets describing sensor properties and process conditions from polymer production in a practical use case, to attain improvements in on-line process analysis and off-line sensor property inspection [[Tax01] [KM10] [Bar10] [K+94a] [TD04] [D+07] [Tax13]].
The *State-of-the-art* methods, such as neural networks, support vector machines, k-nearest-neighbor classifier, and NOVCLASS, modified for parallel two-way recognition of critical operating conditions and faulty sensor properties, and trained on three datasets with a rejection rate of 0.00001, achieved 99.9% accuracy. A second feature reduction by Sammon's Mapping combined with boundary settings visualized the trained and tested datasets in a 2-dimensional view (see Chapters 8 and 9, Fig. 87) in a first step with Matlab and QuickCog.
Further developments should lead to state memory approaches, recording the so far covered distance and calculating 2-dimensional directions, described by a variety of additional attributes.The recognition of novel incidents in advance of occurence could be improved.
The industrial requirements for the recognition of specific faulty sensor properties and critical operating conditions were completely fulfilled. An advanced process control prototype for partially holistic detection shall be implemented based on these observations.

7.2 Trajectory Prediction and Optimization

Today's *OCC and Novelty Detection* approaches analyze single condition points to determine whether the production process behavior has changed from normal to abnormal. Such changes happen due to material property modifications basically caused by environmental pollutants, such as humidity or dust particles, occurring after 10, 20, or 100 min, or never. This implies a short reaction time, within the 10 min following recognition, before the abnormal process deviation leads to material waste in the polymer production cycle. However, a robust polymer process that has already drifted into an abnormal condition point is hard to recover and is usually aligned to a machine breakdown (initiated by the machine or manually).
There are novel approaches for instantaneous value (state) to trend memory classification that deal with these issues of earlier process recognition, supported by interacting dependencies between several condition points in sequence, called **trajectory**. The medical or automobile industries use trajectory functionalities to recognize patterns and create simulations for, e.g., computer-aided robotics and controllers. However, in the polymer film industry there is a lack of trajectory research approaches applying such technology for monitoring and forecast improvements [[W+94] [HvS00] [L+08] [V+02]].

In the following, the previously analyzed *OCC and Novelty Detection* approach will be extended by a *Trajectory Visualization* approach, for advanced recognition and recommendation.
State-of-the-art classification approaches, e.g., neural networks with regard to multi-class classification, have been investigated with existing datasets and previously found results, in order to improve the system reliability for viable online process analysis to achieve best results.
Seven different methods (neural networks, support vector machines, *k*-nearest-neighbor classifier, naive bayes, automated network search, boosted trees, and NOVCLASS) were trained off-line on two datasets, and exemplary results were visualized with Matlab and QuickCog in a first step, as described in the following subsection.

7.2.1 Data Acquisition and Processing

A robust process monitoring quality prediction loop (Fig. 60) based on the previously extracted OCC system sensor machine data, framed *OCC Results* from Data Acquisition, Training, Validation, Testing, OCC Prototype, and Novelty Detection, and extended by the framed *Trajectory Approach* was investigated off-line with supervised methods for Dataset 1, the pressure data.

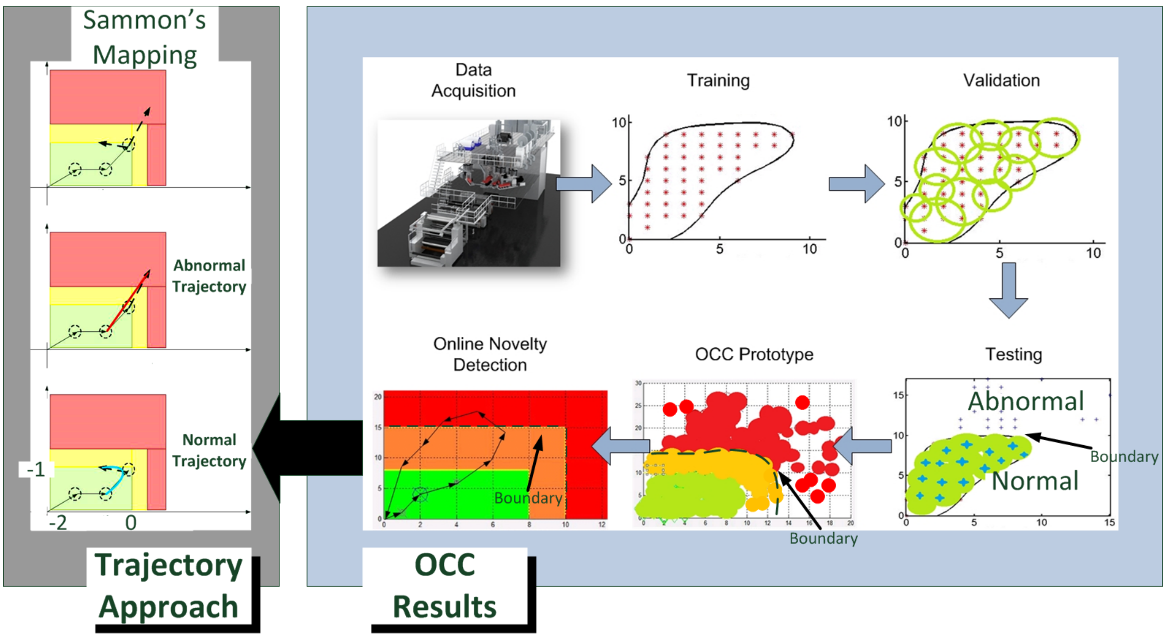


Figure 60 Extended process diagram for data acquisition, training, validation, and testing, developed OCC preliminary prototype, and simplified visualization with process trajectory prediction

In detail, the previously acquired 21,900 datasets – mostly pressure, temperature, speed, consumption, and dosing data – from one extrusion machine were trained (70%) by OCC methods, validated (15%) and tested (15%) after best adjustment, to find the best separation boundaries of the target *green field* and the abnormal *red field* class in the cut *OCC Prototype* of the 2-dimensionally scaled view extracted from the previous multi-dimensional scaling step. All following investigations according to trajectories

were based on this 2-dimensional view.
The Figure 61, extracted from the extended process diagram in Fig. 60, shows the sample trajectories for normal and abnormal behavior with regard to the pressure data.

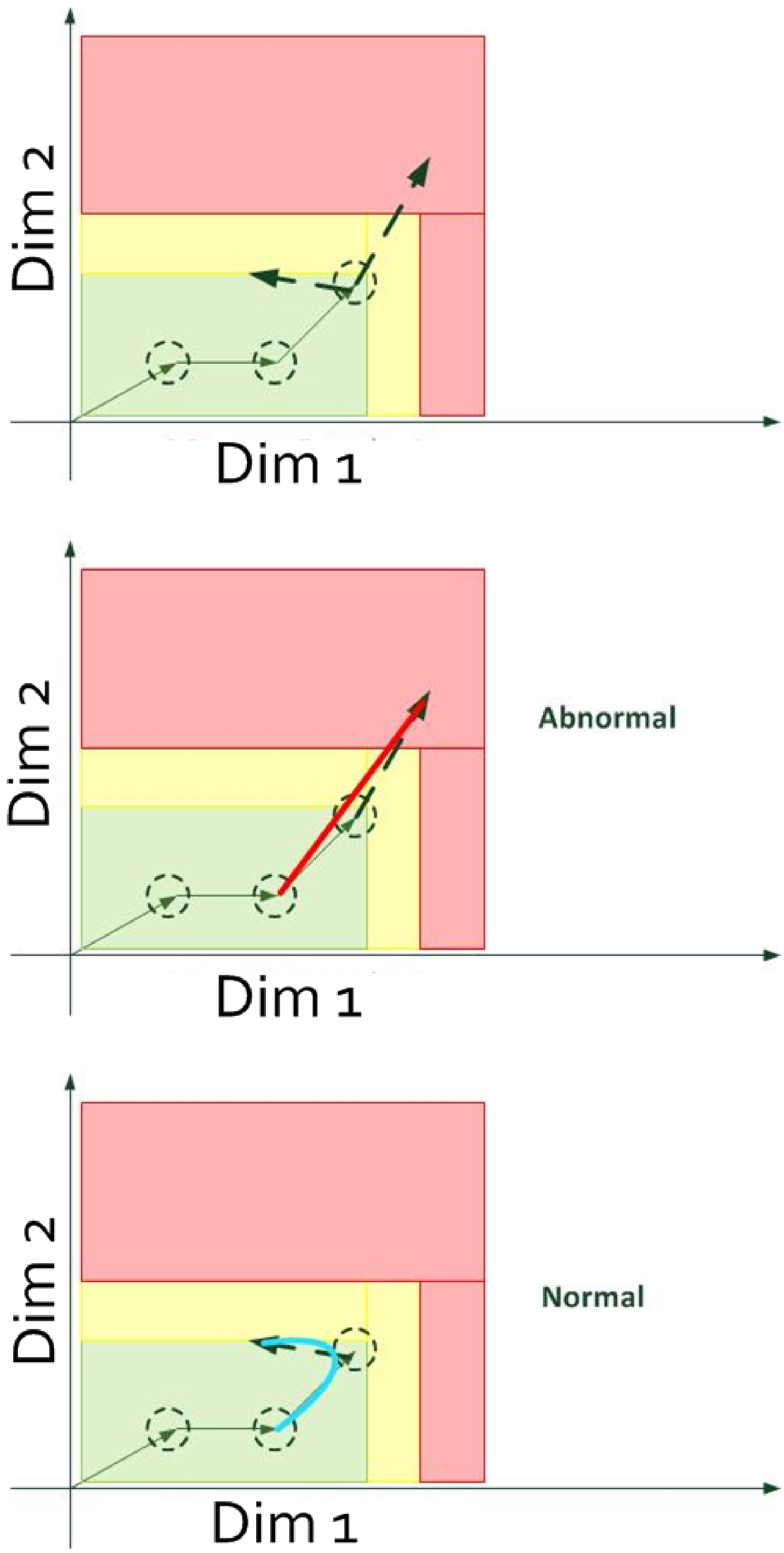


Figure 61 Sample trajectories for possible behaviour by object (top), abnormal behaviour by object (middle) and normal behaviour by object (bottom) based on the previous extended process diagram

The dimensions were visualized for interactive analysis methods with STATISTICA, Matlab or QuickCog. The results from previous investigations were tested on-line afterwards at a prototype machine on a defined waste material "cut-off"' problem (see Chapter 8). Newly extracted observations (Table 19) were separated into abnormal conditions with about 5,500 datasets and normal conditions with about 14,000 datasets, acquired from March to December 2014 for a downstream off-line trajectory analysis. Each final condition recognized as Abnormal Condition = "Shut-Down" and Normal

Table 19 Datasets from extrusion process

Dataset	Type	Dataset Count	Dim	Description
Dataset 4 (Normal)	Extrusion Raw Data	14,000	6	Without faulty data (normal data)
Dataset 5 (Abnormal)	Extrusion Raw Data	5,500	6	With faulty data (abnormal data)

Condition = "Accomplished Product" had a preceding sequence of about 12–199 normal condition points (minutes = timestamps), in sum a so-called trajectory or state memory as a kind of trend analysis, to be explained below as a *Trajectory Approach*.
Compared with the previously explained *OCC Approach*, the novelty detection surveyed single scalar condition points or states which could change. However, during the investigation, the observation of several condition points in sequence offered new insights for a future dynamical predictive recognition of condition points: the trajectory. The more trajectories are investigated, the more precisely an abnormal behavior can be recognized.

Trajectory
A trajectory describes in a 2-dimensional view the condition change over time, beginning at the start of the manufacturing process/run, (Normal Way) moving on in a radial cycle back to the starting point, or (Abnormal Way) in a steep gradient up, and then back to the starting point, as shown in Fig. 62. Each way consists of more points than presented.

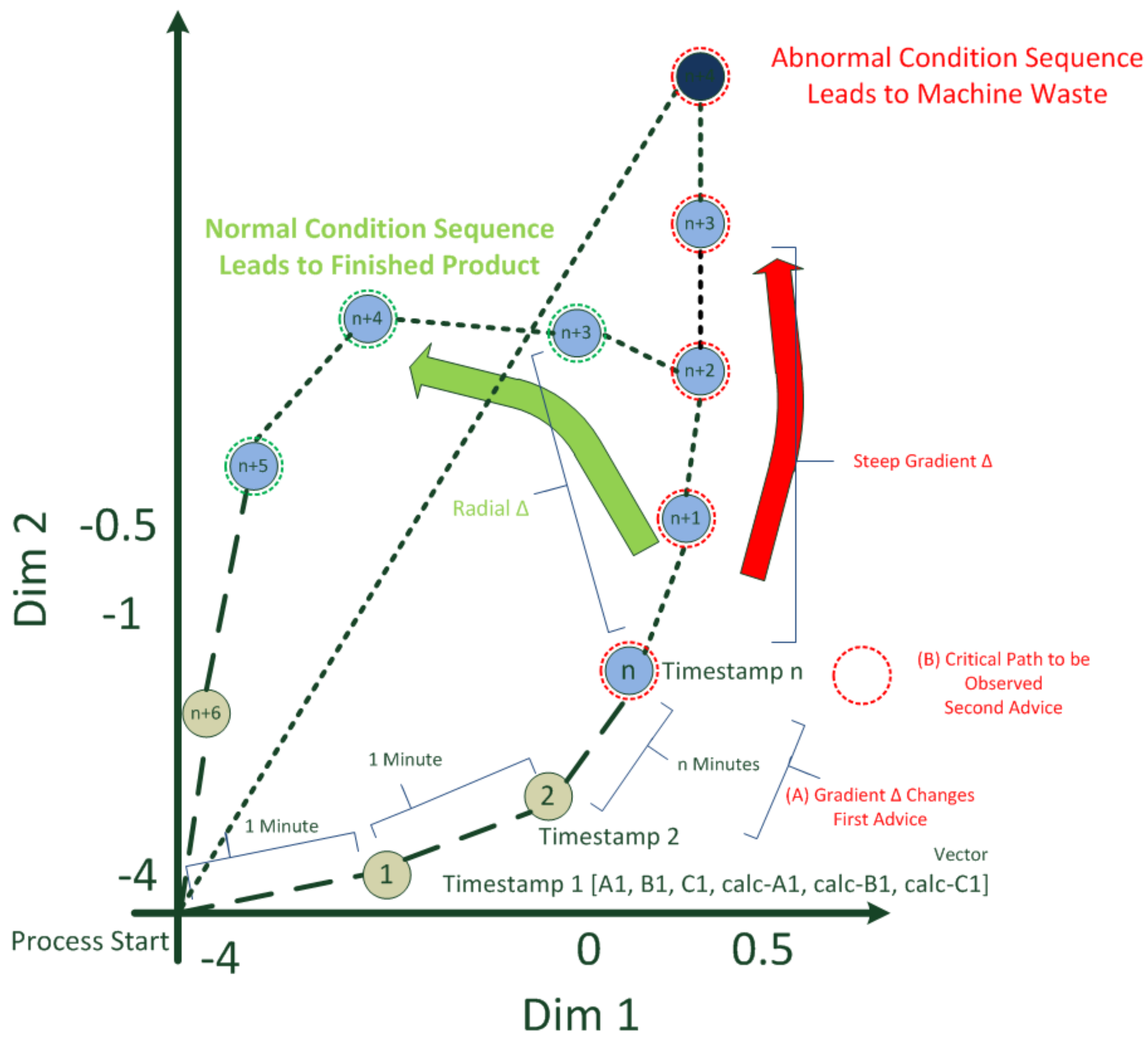


Figure 62 Representative normal and abnormal trajectory course from the process starting point, marked by definite timestamps for typical 120 minutes extrusion product

One trajectory consists of 12–200 acquired sensory datasets, so-called timestamps, displayed in the consecutive sequence as 1, 2, ..., n, n + 1, and so on. Each timestamp includes a multi-dimensional vector of 6 attributes (pressure_A1, pressure_B1, pressure_C1, pressure-calc_A1, pressure-calc_B1, and pressure-calc_C1), in the following named A1, B1, C1, calc-A1, calc-B1, and calc-C1, which are adopted from the previous novelty detection approach.

In Figure 62, the condition point sequence from the new pressure dataset is 2-dimensionally scattered. For further investigation of trajectories, considering 6 attributes at each timestamp but different trajectory lengths, a standard number of timestamps had to be defined to reduce all trajectories to the same length. Therefore, only those parts (timestamps) of each trajectory were investigated that were located within the so-called boundary zone.

In Figures 64 and 65, the processing cycles for recognition of normal and abnormal trajectories, respectively, and their reduction to 12 timestamps is displayed, based on the previous 2-dimensionally scaled plot from Sammon's Mapping.

The complete cycle is described in 5 steps:

(1) The trajectory (sequence of timestamps) starts in the normal zone, then moves to the boundary zone or stays within the normal zone. These are the only possible ways.

(2) The trajectory enters the region of interest, i.e. the boundary zone, by crossing the threshold line. The area is covered by equal-sized squares (Fig. 63) in connection with sliding windows or segmentation approaches [[K+93][DW92][C+04]].

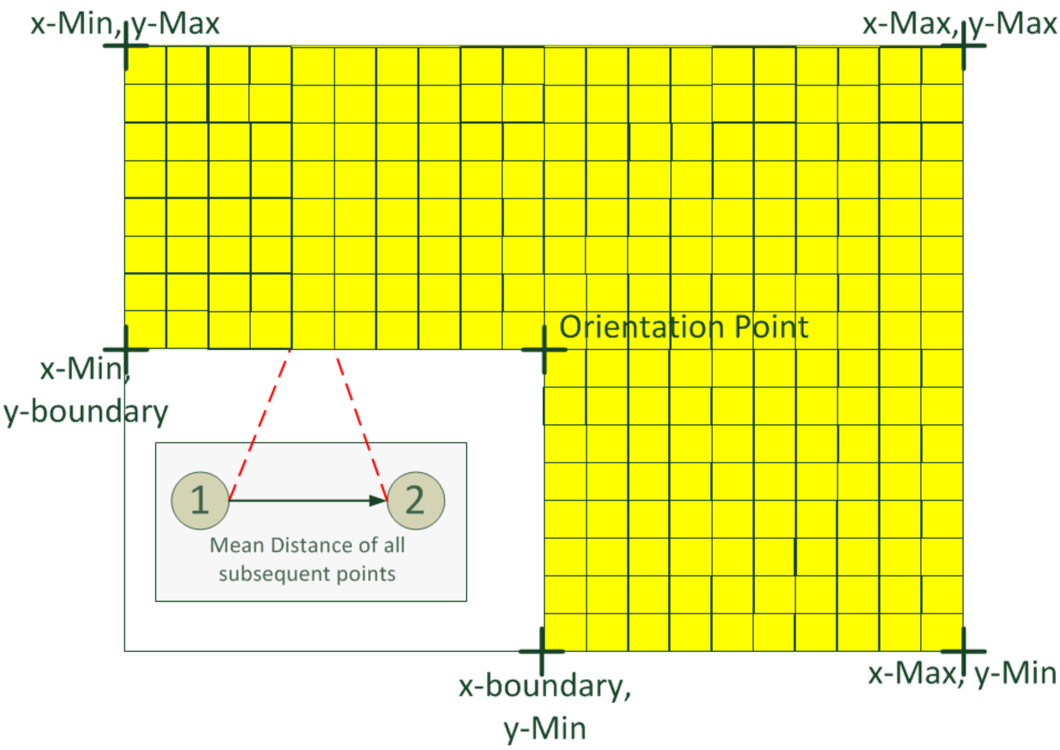


Figure 63 Boundary zone limitations with equal sized squares depending on the mean distance from all regarded subsequent timestamps

The squares were generated as heuristic approach on the 2-dimensionally scaled view by taking the mean distance between all regarded trajectory points (Fig. 63), which is an approximation to the 6-dimensional space. The zone limitations concurred with the maximum and minimum of the training datasets from the Sammon's mapped final plot.
(3) Each time a trajectory enters the boundary zone, a counter begins to count the numbers of crossed squares within the boundary zone. If the trajectory enters and leaves the boundary zone before a fixed number of squares are counted, the count is set back to 0, starting with a new entry. The optimized fixed number (12 in this case) is searched experimentally.
If the trajectory jumps between two squares, it is not recognized as a trajectory for as long the count of 12 is not reached. The trajectory acquisition is finished after 12 squares are counted, each square consisting of 1–30 timestamps. The maximum count of 12 was chosen according to the minimum crossed squares from all trajectories.
(4) The count of timestamps per square is reduced[31] to the first acquired timestamp per square, with regard to similarity by their close location.
(5) The trajectory sequence is defined by the resulting timestamps in the order of their appearance. The attributes are also merged in the order of their appearance.

[31] Smoothing

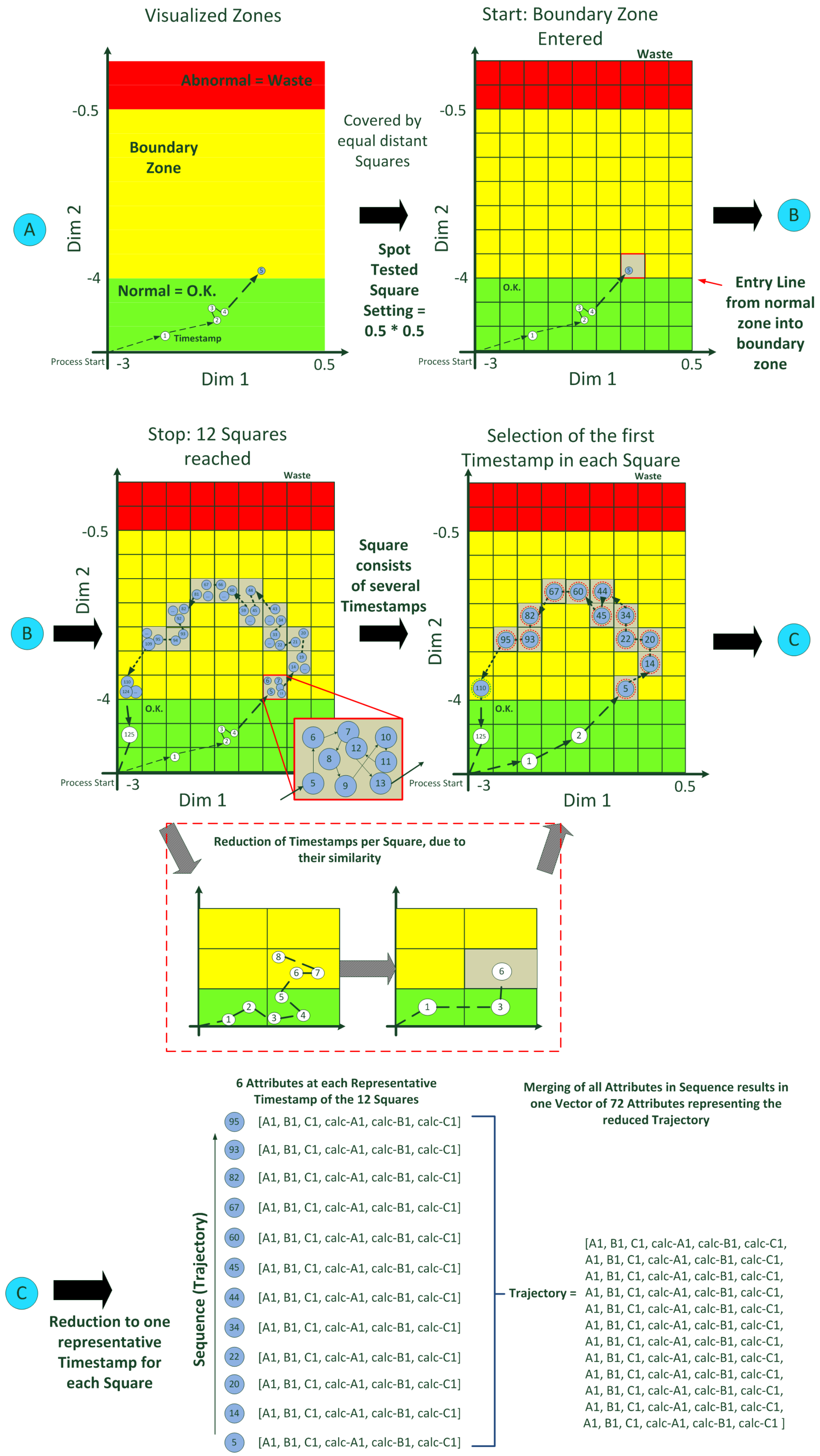


Figure 64 Representative normal trajectory recognition and reduction process: (1) visualized zones, (2) square covered boundary zone entered, (3) 12 square limit reached, (4) timestamp reduction within each square

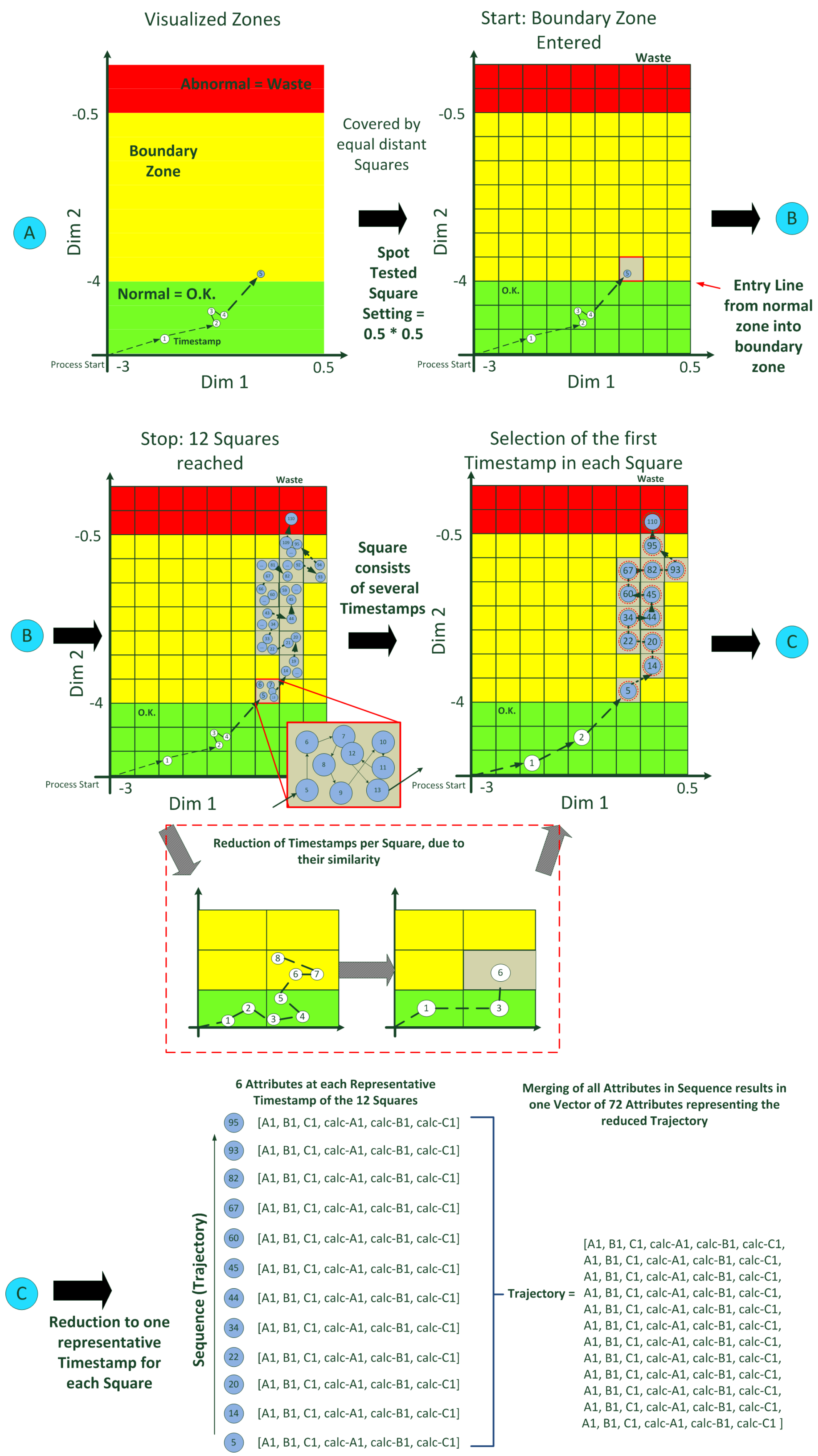


Figure 65 Representative abnormal trajectory recognition and reduction process: (1) visualized zones, (2) square covered boundary zone entered, (3) 12 square limit reached, (4) timestamp reduction within each square

In this case, only parts of the trajectories going through the boundary zone, upwards into the abnormal zone, or in a radial left turn back to the normal zone were investigated. The normal and abnormal trajectory process from the point of entering into the boundary zone, the reduction to significant representatives, and the merging into a vector is shown in the above-mentioned figures.
The typical approach used by optical control systems to define geometrical attributes from the 2-dimensional point of view, like angle values, shape factors, X-Y positions, roundness, length, width, radial or gradient functions for the different trajectory types, was not pursued. Instead, the attributes from the 6-dimensional dataset per timestamp were more robust for use in case that in future approaches the 2-dimensional view should change to other scatter attribute views.
Regarding a possible 12–199-min trajectory time range, different trajectory lengths resulted from the varying process speed between 10–150 m/min, manually adjusted by the staff. For all trajectory lengths, the timestamp range chosen by spot test of 30–60 min, i.e. the boundary zone, was investigated due to its location in the previously described yellow field in Fig. 57) (Subsection 7.1.4). The trajectory starts from the initial process point, takes its way over 1 and 2 timestamps and reaches the *yellow field or boundary zone* at n timestamps. Different trajectory lengths exist for abnormal conditions, e.g., one consisted of 40 timestamps and another one of 50 timestamps. Therefore, no equal numbers of timestamps and same numbers of attributes that have to be merged afterwards exist for distance calculation. In order to achieve that both trajectories are displayed by 12 instead of 40 or 50 timestamps, equally distant squares were used within the boundary zone, distinguished into Dimension 1 and 2 directions. The **square size** of 0.05 * 0.05 (Dimension 1 * Dimension 2) was chosen by spot test in the 2-dimensional view. The size is defined as the mean distance between timestamps of all regarded trajectories on the scaled plot. A so-called equal-distance square matrix covered the 2-dimensional view. As soon as 12 squares were crossed in the described sequence, the timestamps were stored and reduced within the square to one representative each, and the trajectory is defined by successive 1 – 2 – ... – 12 timestamps for classification. Thus, different trajectories are standardized to a uniform length.
After level 12 with an average time range of 30–60 min (timestamps) in advance of an occurrence, predictions on the directions and sizes are possible, for the normal or abnormal final conditions.
Each timestamp consists of a 6-dimensional vector with pressure attributes. In the last step of the trajectory definition, the 12 timestamps with 6 attributes were merged to a trajectory of 72 attributes (dimensions), presented in Fig. 66 according to Figures 64 and 65. For a better understanding, an additional graphic of the merging of 4 attribute vectors is shown.

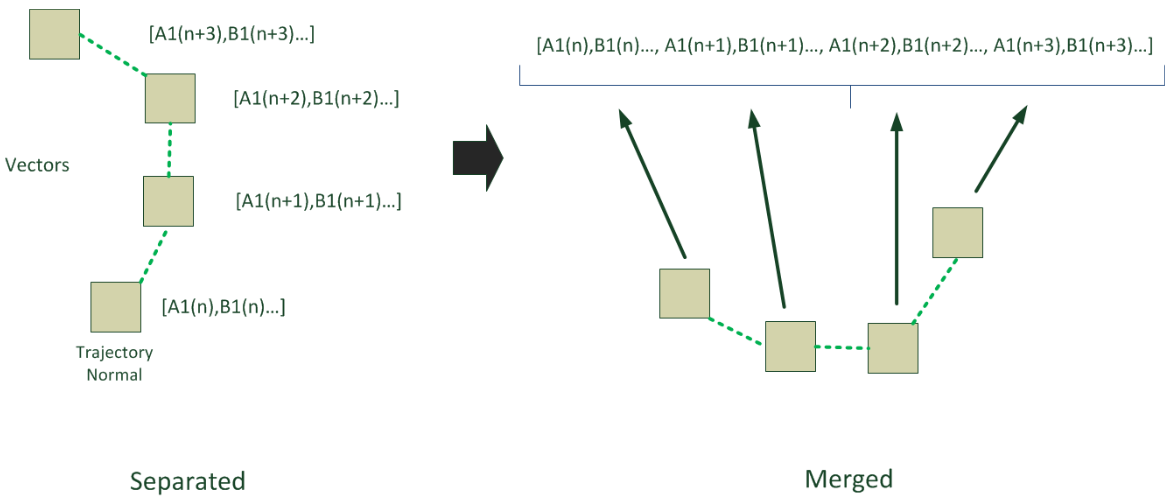


Figure 66 Exemplary: separated 4 vector attributes merged to one 24 attribute vector for metric distance calculations

For distance calculations, the vector attributes from each level were merged to 1 vector, handling each timestamp attribute, e.g., A1(n) or A1(n + 1), as two unique attributes. Thus, for a 12-level case, 1 vector of 72 attributes is merged starting from 12 vectors with 6 attributes each.
Thus, 257 trajectories of the same length were generated from the 19,500 new pressure datasets, separated into 120 normal and 137 abnormal trajectories (Table 20).

Table 20 Datasets from extrusion process

Dataset	Type	Trajectory Count	Dim	Description
Dataset 4 (Normal)	Extrusion Raw Data	120	72	Without faulty data (normal data)
Dataset 5 (Abnormal)	Extrusion Raw Data	137	72	With faulty data (abnormal data)

Model Building Part

The classification of new observations depends on the classification metrics used, such as the Euclidean, Manhattan, or Mahalanobis Distance for an exemplary k-nearest-neighbor classifier (Subsection 4.2.4). To perform the allocation of new observations to the classes (normal and abnormal), the – in this case – 72-dimensional Euclidean Distance metric is calculated for each new trajectory object. In an exemplary k = 3 nearest-neighbor classifier case, the nearest 3 trajectory neighbors within the 72-dimensional space are regarded. The nearest calculated distance to the majority of nearest-neighbor trajectories defines the associated class for the new observed trajectory. In the following investigations, the 72-dimensional vector attribute space was taken as a basis for calculations on the previous defined datasets.
The trajectories generated from the Datasets 4 and 5, defined in Table 19 and Table 20 for classification and extracted from the previous novelty classification with a higher amount of timestamps, are considered in the following as they offer potential for supervised learning methods.
The trajectory approach is an extension (downstream classification) of the pre-stage novelty detection, as presented in Fig. 67.

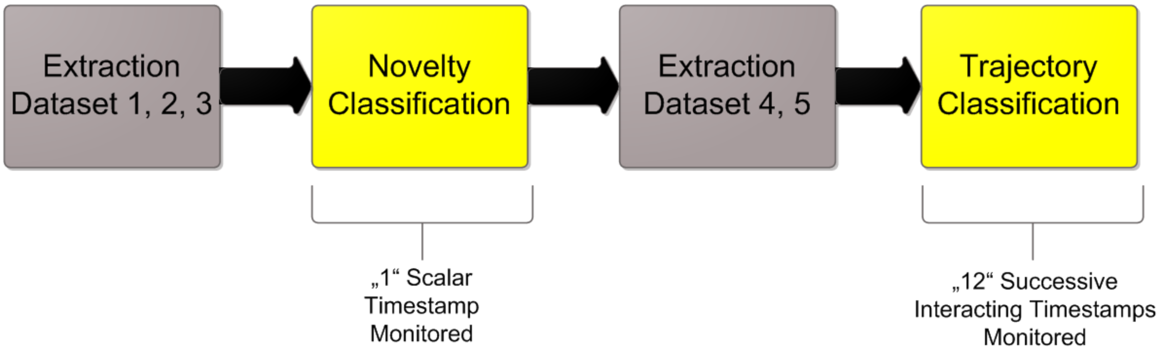


Figure 67 Process diagramm for data acquisition, training, testing, developed (trajectory) preliminary prototype, and final trajectory classification

The cycle from data extraction, novelty classification with regard to one scalar timestamp for condition changes, and the generation of new datasets out of it, to the second classification stage regarding successive interacting timestamps is a novel approach for the polymer film industry. The results are presented in Subsection 7.2.2.

7.2.2 Results Examination

For objective decisions on future process conditions, the operators' experience, laboratory quality measurement systems, and inline control systems are *State-of-the-art* for quality information in production processes and thus do not efficiently predict critical states in time.
The novelty detection methods can be optimized by trajectory classification, which allows a finer configuration and lowers the uncertainty of recognition by examination of parameters and methodological settings.
The final process trajectory presented in Fig. 68 is in the focus of the investigation for faster prediction of faulty process conditions by supervised learning methods. Therefore, the defined trajectory timestamps are used for feature computation.

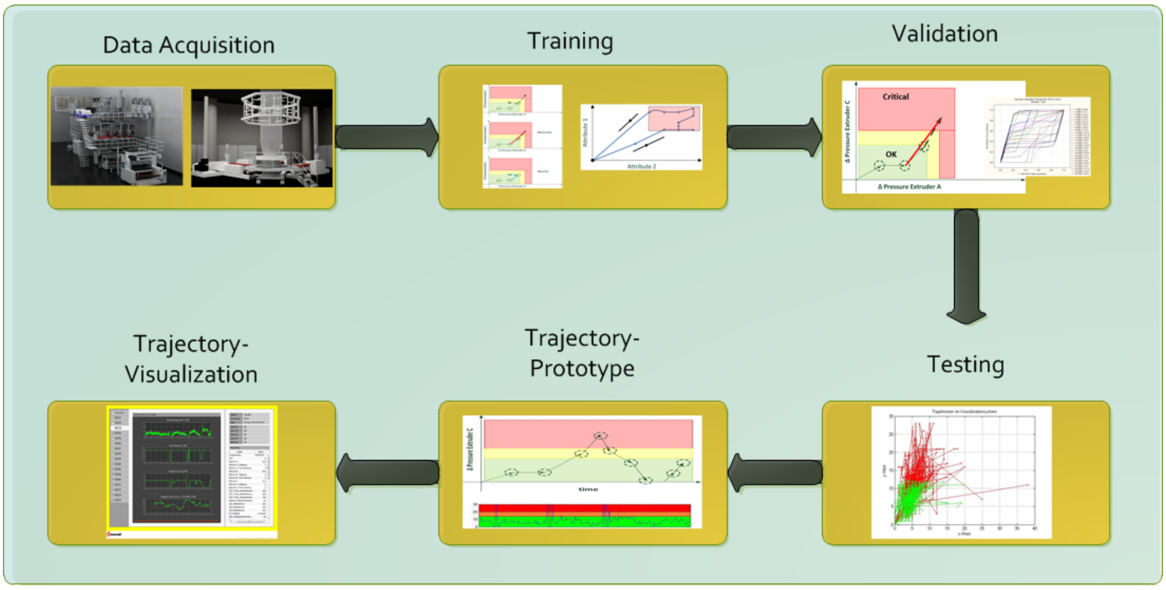


Figure 68 Process diagramm for data acquisition, training, testing, developed (trajectory) preliminary prototype, and final trajectory visualization

The complete process cycle consists of the parts, data acquisition, training (70%), validation (15%), testing (15%), developed preliminary prototype, and conceptual *Trajectory-Visualization*, regarded in Chapter 8. In this Subsection, the preselected methods are experimentally adjusted to achieve best results for the two-class classification on two defined datasets from a real production process. As described in Subsection 5.2.1,

Table 21 Sensitivity Analysis for general methods

Methods	Classifier Type	Best Settings
Neural Networks	RBF	HiddenNeurons = 8; Rej = 0.00001
Support Vector Machines	RBF	Sigma = 5 and 10; Gamma = 0.1; Rej = 0.001
k-Nearest-Neighbor Classifier	–	k = 5; Rej = 0.001
Naïve Bayes	–	Probability = 0.32/0.67
Automated Network Search	MLP	42-23-2
Boosted Trees	–	Optimal number of trees: 186; Maximum tree size: 3
NOVCLASS	–	ScaleFactor = 1.7 ;

the dataset selection (hold-out, 70% training, 15% validation, 15% testing) consists of pressure, temperature, speed, consumption, and dosing data chosen based on the operators' experience and quality measurement systems. The settings as described in Table 8 were modified consecutively (e.g., k-nearest-neighbor classifier: $k = 1$, Rej = 0.1; 0.01; .../$k = 2$, Rej = 0.1; 0.01; ...), and the empirical lowest repeatedly received ACC results after ten runs (STD < 2.5%) were kept or refused depending on a suggested significance test (t-test, p value 0.05) and best settings as displayed in Table 21.
The methods (neural networks, support vector machines, k-nearest-neighbor classifier, naïve Bayes, and automated network search were investigated off-line. With regard to the *OCC/ Novelty Detection*, the classifier methods were extended by naïve Bayes, automated network search, and boosted trees, and Datasets 4 and 5 were extracted from the previous novelty investigation. Due to the sparse outlier samples, the OCC accuracy improvement was focused on the normal *One-Class* datasets. In the following trajectory analysis, the normal Dataset 4 (trajectory count 120 from 14,000 datasets) and the abnormal Dataset 5 (trajectory count 137 from 5,500 datasets), associated in a time chain with the previous process states novelty detection, were both considered as two classes with similar amounts of datasets. In this case, two classes with different conditions (normal and abnormal) with equal amounts of datasets were analyzed, with the above listed *State-of-the-art* classifier methods, regarding their accuracy ratio and experimental standard deviation, which are based on the previously found OCC Selection and Evaluation results.
The investigated features deduce machine properties (EP-Feature) from the extrusion stage. The best settings for the normal and abnormal trajectory directions, displayed as green (left direction) and red (upward direction) part within the visualization from Fig. 61, were achieved by the boosted trees method (number of trees = 186) and, for the examined data (tree size = 3), are presented in Table 22. Boosted trees achieved the best specificity rate for Dataset 1 with 100% for normal datasets and 99.6% for abnormal datasets. The best fitting parameter settings according to the results are displayed in Fig. 21. The accuracy for normal datasets was low due to the different

Table 22 Experiments' Accuracy with 7 methods, and 2 datasets (normal and abnormal datasets used)

	Methods	Dataset	Condition Class	Mean in [%]	STD in [%]	AbsErr [#]
1	Neural Networks	Dataset4	Normal	88.4	0.2	14
	(RBF)	Dataset5	Abnormal	92.2	0.2	11
2	Support Vector	Dataset4	Normal	84.0	0.2	19
	Machines (RBF)	Dataset5	Abnormal	94.6	0.6	7
3	k-Nearest-Neighbor	Dataset4	Normal	80.4	1.6	24
	Classifier ($k = 5$)	Dataset5	Abnormal	94.0	2.0	8
4	Naïve Bayes	Dataset4	Normal	61.2	2.3	47
		Dataset5	Abnormal	94.1	2.0	8
5	Autom. Network	Dataset4	Normal	95.8	0.2	5
	Search (MLP)	Dataset5	Abnormal	98.2	0.1	2
6	Boosted Trees	Dataset4	Normal	100	0.0	0
		Dataset5	Abnormal	99.7	0.7	1
7	NOVCLASS	Dataset4	Normal	92.2	1.2	9
		Dataset5	Abnormal	97.1	1.6	4

types of trajectories for normal conditions and their starting behavior, which is very similar to the abnormal datasets; however, the abnormal dataset results were in the focus of this case. All repeatedly achieved mean results after ten runs were monitored. The accuracy rate defines the error in the target class, i.e. the possibility of becoming classified to the wrong class.

Automated network search (MLP), boosted trees, and NOVCLASS achieved more than 90% accuracy for abnormal datasets, which corresponds to the accuracy of optical control systems for visual defect detection. The boosted trees, in comparison to neural networks, run efficiently on large datasets, are robust to noise and can be extended to unsupervised unlabeled data approaches. Mathematically, it is a gradient method, dividing the set of data into subsets best fitting to the target class with a low amount of settings.

One approach for off-line analysis of the quality of process sensor property prediction with a trajectory classification system based on a previous OCC analysis of machine sensor data was improved by additional classification methods. After best adjustment of the classifiers to the datasets, the methods should be implemented into a novel design graphical user interface (GUI), with e.g., Matlab & Simulink, supporting the online process control by the operators, besides optical control systems and laboratory results.

Figure 69 Novel design graphical user interface, developed with Matlab & Simulink

Such a prototype system (Fig. 69) has to be tested on interface settings, recommendation tasks, and visualization parts, as introduced in Chapter 8.

7.2.3 Complete Assessment and Modification Loop

In summary the significant processing steps assigned to the classification methods and to the staff visualization part are presented the following.

Staff Visualization Cycle:

- (1) OCC, Novelty Classification
- (2) Multidimensional Scaling – 2 D - Sammon's Mapping
- (3) Maximum Setting
- (4) Interface Implementation - Scatter Plot, Traffic Light (Chapter 8)

Classification Approach:

- (1) OCC, Novelty Classification
- (2) Multidimensional Scaling – 2 D - Sammon's Mapping
- (3) Maximum Setting
- (4) Trajectory Reduction
- (5) Trajectory Classification

The following Pseudo-Code 11 is connected to the previous *OCC Selection and Evaluation* 10, and summarizes the described *Trajectory Visualization* in subsequent steps one by one:

begin

8. **Projection of New Testing Data into 2D Mapping**
   Sammon's Mapping Recall [Kö00]
   Projection of new Data-Points into previous 2-D Mapping
9. **Boundary Zone Definition**
   Framework
   Boundary Line set by 30% in Dim1 and Dim2 direction
   Covered Squares starting at [-4, -4] size 0.5 * 0.5
10. **Trajectory Acquisition**
    Crossing Boundary Line Counter Start 1:12
    if square is entered, store current timestamp
    end if 12 stored timestamps reached
    Reset Counter to 0 if 12 reached or zone was left
11. **Vector Merging**
    Combining Attributes 1 Vector = 6 Dimensions
    Adding all Vectors in Sequence of Occurence
    Trajectory Attributes A = 12 Timestamps * 6 Dimensions
12. **Model building with Classifier Methods**
    k-nearest-neighbor
    Cross Validation of 257 Trajectories into 70%/ 15%/ 15%
    Supervised Learning with k-nearest-neighbor
    k = '5', rule = 'nearest'

end

Pseudo-Code 11: Procedure of the OCC selection, evaluation, scaling and projection approach

7.2.4 Summary

Using *State-of-the-art* classification methods for the trajectory prediction of critical process conditions in polymer film production is a novel approach in the rigid-film industry. The trajectory approach is based on datasets from previous OCC results. Normal and abnormal condition states of process behavior were investigated for separation into two classes. *State-of-the-art* classification methods were examined on predefined datasets, characterized by their time shift. The classification result of 99.6% accuracy for abnormal conditions supports early novelty recognition, so that the method will be integrated into standalone applications, as described in Chapters 8 and 9.

Further investigations should focus on the diagnostic cycle of prediction and real incident, which is currently manually accomplished by comparison of the waste rolls to their classification by the staff. A state memory could increase the stability of prediction.

The recognition results from the polymer production process can be used for high-level failure prevention of the complete supply chain when transferred to other objects such as the process lines, sub-divisions, divisions, plants, corporate group, or the market.

7.3 Prediction and Reduction of Energy Consumption

In addition to trajectory visualization, the energy trends were investigated as a second research vehicle, with regard to the prediction of the end year energy hours measured as 7,000 h. With about 220 measurement sensors, e.g., Janitza[32] UMG, the energy acquisition network was established at the Mondi Gronau GmbH from 2012 to 2015. Measured media types were electricity, gas, and water. A *EN ISO 50001*-certified plant-integrated energy management system monthly reports the datasets acquired internally and externally within 15-min intervals to the suppliers, EnBW and Stadtwerke Gronau. The main electrical sensor data from 2001 to 2014 characterizing the whole plant were investigated on a monthly base. Each yearly trajectory consisted of 7 timestamps, from January to July, with 3 attributes: *Peak Load* in [kW], *Consumption* in [kWh], and *Proportion* in [kW]/[kWh]). Two-class classification methods were examined to find the best fitting method for energy hours higher than 7,000 h or lower than 7,000 h. In Fig. 70, the net fee probability is displayed, depending on the prediction point, the monthly basis in this case.

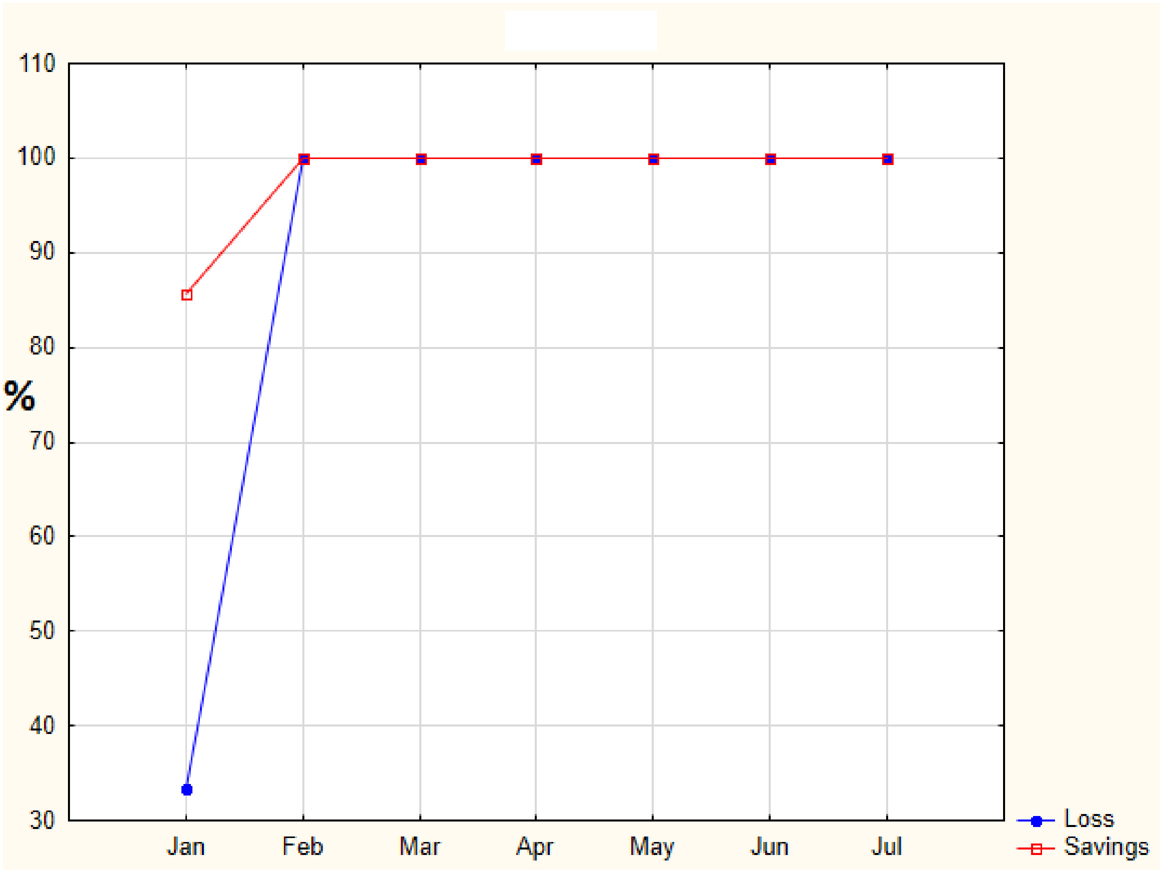


Figure 70 Energy net fee prediction probability [%] from January to July 2014 for fee loss and savings (MLP, 12-25-2)

The best prediction probability of 100% (Fig. 70) and 0.3% experimental standard deviation in ten runs was achieved with a random chosen method out of a bunch of

[32] Supplier of Measurement Systems

further investigation methods, the multilayer perceptron (12-25-2) in STATISTICA. The deposited plant energy datasets from 2001 to 2014 consisting of two classes (*Loss = Net Fee Payed* \ *Savings = Net Fee Not Payed*) were randomly selected by the hold-out method and afterwards divided into a training set (70%) and the rest, followed by a second hold-out selection into the validation (15%) and testing datasets (15%).
The off-line examined datasets from 14 years of energy consumption at Mondi Gronau GmbH were analyzed and stepwise modified until best results of more than 90% accuracy were achieved with different *State-of-the-art* classifier types.
The predicted result of energy hours lower than 7,000 h was approved by a separate calculation of 6,900 h.
Afterwards, the results were compared to the net stability ratio from EnBW: 6,895 h were reached. Therefore, the net fee of about EUR 800,000.00 had to be paid.
The controlling and energy management teams benefit from an early prediction of the energy consumption behavior to initiate preventive actions, e.g., lowering the load peaks.

7.4 Interview and Operator Feedback

According to Chapter 6, *(B) Monitoring & Recommendation*, different groups of employees were interviewed, and a first version of a prototype monitoring interface with recommendation add-on was introduced, presented in Table 23. The first operator

Table 23 First version extraction of the multiple response discussed interface settings: approval and refusal

Steps	Property	# 1	Appr.	Ratio in [%]	# 2	Appr.	Ratio in [%]
1	Message System	Email	✓	40	Report		0
		SMS		0	Webpage		0
		Traffic-Light		0	Dashboard		0
		Call		5	Application	✓	70
2	2-D Plots	Zoom		20	Refresh. Rate	✓	0
		Drag & Drop		0	Fixed	✓	0
		Limits		70	Variable		0
		Sizable		0	Replay Mode		0
3	Levels of Recom.	1 – 3		0	Coloured	✓	60
		4 – 5	✓	100	Message Box	✓	100
		6 – 7		10	German	✓	100
		8 – 10		10	English		10
4	Interface Location	Terminal		0	Home		0
		Office	✓	90	Entrance		0

feedback on the prototype monitoring & recommendation interface achieved 50% of favored implementation parts, as finally established in the following. In about ten further meeting steps, an assessment type according to [Dex06] was developed, and the software concept and investigation results were discussed with 20 internal employees

(machine operators, shift supervisors, application engineers, division managers) and 3 external software companies in an open, non-standardized (open, oral, written, expert and narrative) questionnaire, with multiple markings admitted. Besides presenting own ideas, each person was able to comment on other concepts with three response types: yes, no, no comment. Each survey was recorded in a separate log file for further investigation.
The 40 prioritized basic information batches regarded were improved in the subsequent interviews, e.g., which attributes from pressure, temperature, speed, and their deviations, variances, or calculated attributes to be monitored. Each participant prepared a list of about 30–300 important attributes; the attributes were merged and reduced in the end to one set per machine. In Table 24, the four mainly requested software property steps and their interview ratios are displayed. According to the interview, the

Table 24 Final version extraction of the multiple respone discussed interface settings: approval and refusal

Steps	Property	# 1	Appr.	Ratio in [%]	# 2	Appr.	Ratio in [%]
1	Message	Email	✓	90	Report	✓	50
	System	SMS	✓	80	Webpage		40
		Traffic-Light	✓	90	Dashboard		40
		Call		5	Application	✓	90
2	2-D Plots	Zoom	✓	60	Refresh. Rate	✓	100
		Drag & Drop	✓	60	Fixed	✓	60
		Limits	✓	100	Variable		30
		Sizable		0	Replay Mode	✓	30
3	Levels of	1 – 3	✓	90	Coloured	✓	80
	Recom.	4 – 5		10	Message Box	✓	80
		6 – 7		0	German	✓	100
		8 – 10		0	English		10
4	Interface	Terminal	✓	80	Home	✓	60
	Location	Office	✓	80	Entrance		0

IT department checked the attributes for static limits, deposited in the database. If no static limits exist, the software should, e.g., automatically calculate dynamical 2.5% standard deviation limits within the given time range.
The interview proposals from 1–10 plot types were reduced to four plots presenting real-time data from the previously selected attributes by drag & drop. Each plot shows time series of 1-min intervals from 1–7 days back, with a refreshing option between 1 and 120 min individually tunable by the user, Fig. 71.

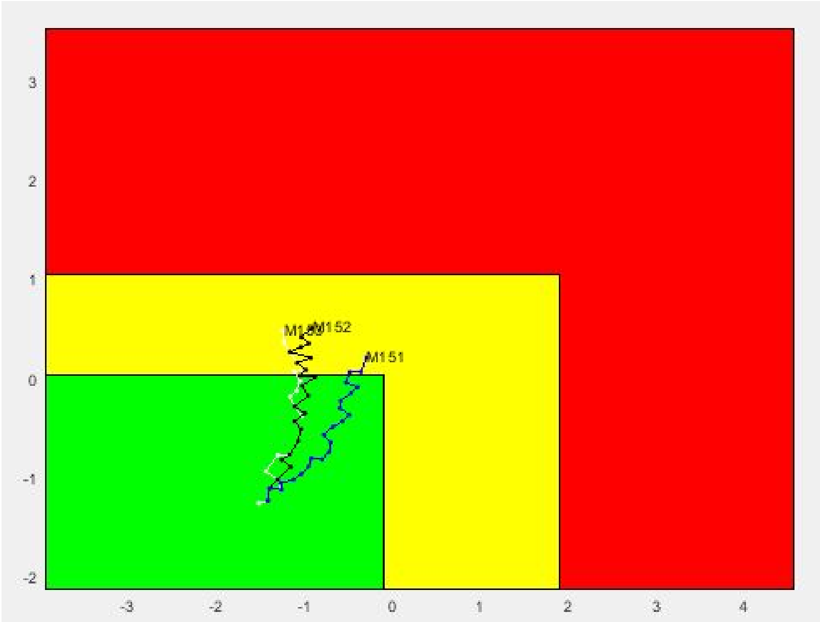


Figure 71 Dynamic 2-D plot real-time trajectory behaviour of different machines

The monitored attribute limits automatically switch from static to dynamic, which is displayed in the plots and in the current value fields.
Additional implemented software features (such as options for saving screenshots or sending e-mails or SMS messages, mobile device usability and fast refreshing rates of < 1 min) were indicated in a feedback form to inform the supervisors about occurring deviations.
The message recommendation part for the interface and the e-mail and SMS transfer, as message box and traffic lights, was reduced from 5 to 3 colored classifier levels.
The demonstrative plot zone description for the staff recommendation was adapted to the machine instruction for waste product rolls and normal product rolls, according to traffic lights in industry and consumer protection:

- Green : Target Zone → Everything is fine
- Yellow : Boundary Zone → Warning, Quality Problems Predicted
- Red : Outlier Zone → Shut-Down and Cleaning after roll is finished

As shown in Fig. 72, the level value was also transferred in a first step to the maintenance panel for giving information about current conditions to the engineers, who are thus able to prepare themselves earlier for services.

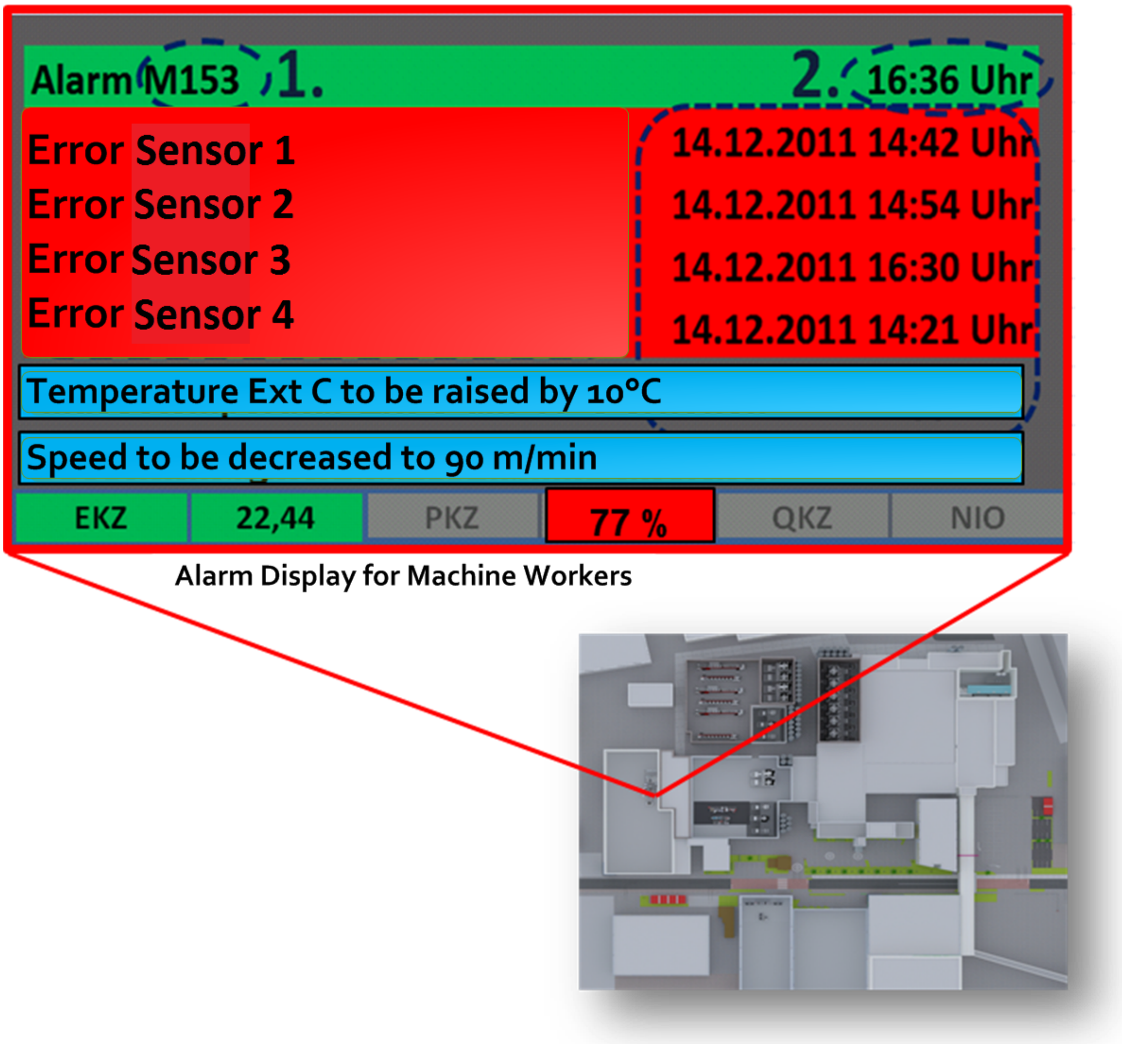


Figure 72 Additional production information interface for recommending parameter settings to the operator, translated into english

Novel ideas were expressed for interface locations; of these, two redundant locations were focused on. The interface is installed on several office computers as a kind of monitoring station, to be also accessible at home. The second, more integrated installation at the machine terminal was mainly preferred, offering local observation possibilities.
The programming and compilation was done with Matlab & Simulink, for reasons of best flexibility in a standalone design compared to STATISTICA, SAS Visual Analytics, Rapidminer, and QuickCog.
For energy visualization, two out of three options were selected. The recommended e-mail energy report and television visualization on a webpage were chosen, whereas the compilation of a standalone application was refused due to higher efforts.
The decisions for all settings are based on the approval of at least 80% interview participants, as previously described.

7.5 Discussion

In a first approach, multi-class classification methods were extended for anomaly recognition of real production datasets. For process state prediction and process yield optimization, different *One-Class* classification methods such as *One-Class* support vector machines, *One-Class* k-nearest-neighbor classifier, *One-Class* neural networks (auto-encoder), and *NOVCLASS* were tested off-line. Although all methods achieved best results, a preselection of three classifier types was chosen for further implementation, due to their simplicity of installation into a Matlab & Simulink environment: the *One-Class* support vector machines, the *One-Class* k-nearest-neighbor classifier, and the NOVCLASS due to its novelty characteristics.

In a second step, seven *State-of-the-art* classification methods (neural networks, support vector machines, k-nearest-neighbor classifier, automated network search, boosted trees, and NOVCLASS) were trained off-line on datasets from the previous investigation; exemplary results were visualized in a first step with STATISTICA, Matlab and QuickCog and extended for energy consumption trajectory investigation. The chosen methods for further implementation were automated network search and boosted trees, due to best classification results for both classes. Additionally, NOVCLASS and the k-nearest-neighbor classifier were picked for further investigations of new datasets in upcoming trajectory research studies.

Three different datasets from the polymer extrusion case study were analyzed; the settings for each method were adapted for best recognition rates and the final implementation was discussed with the participants. After best adjustment of the modified classifier methods to the three datasets for novelty detection, results of 99.9% accuracy were achieved for abnormal conditions. Downstream, a second recognition accuracy of 99.6% for abnormal trajectory behavior was additionally accomplished and continuously tested in a novel approach for energy consumption behavior, with 100% accuracy.

The interview and operator feedback investigated the idea of a monitoring interface in relation to the results found in Chapter 7 by a committee of plant experts, supported by external companies, to assemble a choice of necessary features for standalone applications, including the previously regarded adapted methods. The next step should be the integration of the methods into a closed in-line testing environment for procedure velocity improvement and adjustment of the interacting monitoring & recommendation parts, such as real-time data acquisition, interactive visualization, early recommendation, and realization of the requested interface settings.

Therefore, in the following Chapter 8, the main focus for examination lies on the interface settings and on the implementation cycle within a laboratory prototype testing approach.

8 Experimental Laboratory Monitoring Investigation Approach

In this chapter, the experimental on-line laboratory prototype for real-time monitoring is investigated and demonstrated on the currently implemented system architecture described in Chapter 6.

A graphical user interface, that is novel for the polymer film industry, with near-real-time acquisition, processing, visualization, and recommendation application, a so-called *Laboratory Data Acquisition* ***Con****dition Real-Time* ***Mon****itoring System* (**ConMon**), was established in a preliminary step for a prototype *cast* extrusion machine (see Chapter 3) and optimized, to provide added value for the running production.

The complete setup was built based on one prototype research machine, which was separated from the main production process but connected to the internal machine network (VLAN). The preferred data stream cycle is displayed in Fig. 73,

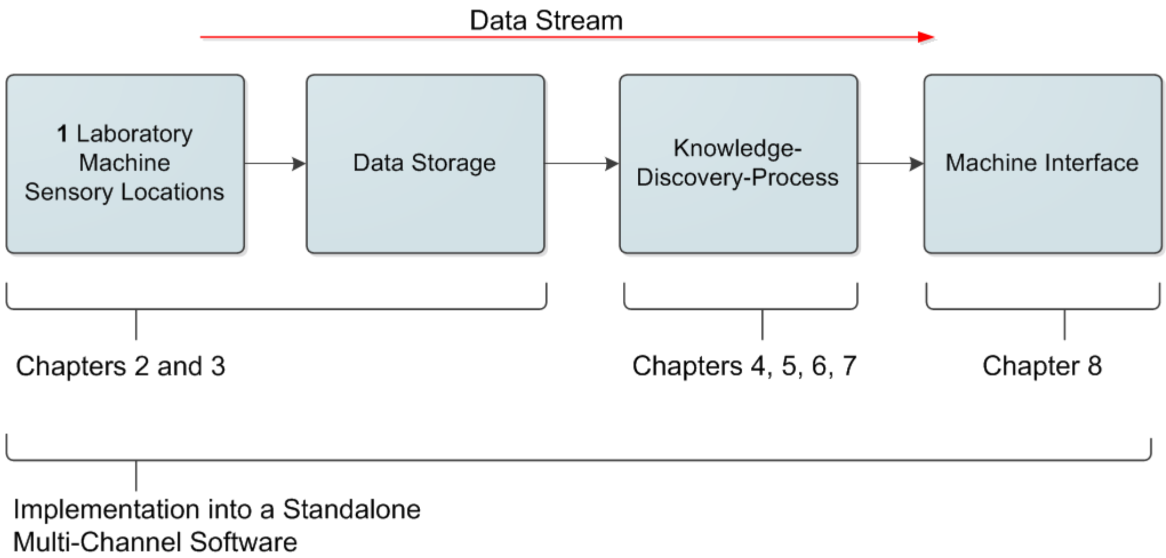


Figure 73 Experimental laboratory data stream concept: machine sensory locations, data storage, *Knowledge-Discovery-Process*, Machine Interface

The complete "Data Stream" from the laboratory machine sensor locations to the interface visualization is divided into three groups. The first group includes the sensor locations, the communication links, the data acquisition, the data types, and the storage settings detailed in Chapters 2 and 3. The second group deals with the dataset processing, feature reduction, and classification described in Chapters 4–6, based on the *Knowledge Discovery Process* introduced in Chapter 2. The third and final group, detailed in Chapter 7, comprises the datasets and methods used, and the result visualization and recommendations according to the staff requests (see Chapter 7). The complete setup was implemented into a real-time monitoring data acquisition & recommendation standalone application.

8.1 Data Acquisition, Processing and Visualization

The *cast* extrusion process machine M150 was selected due to the completely finished network connection and its instable process behavior with regard to a specific waste problem, the so-called “cut-off”. In the main extrusion machine process (see Chapter 3), granular polymer (SBS) from the 5 dosing stages is fed into the extruders (3), with one screw conveyor each, producing three film layers at 180–190°C. The formed film coming out of a wide (3.5 m) nozzle is rolled up and slitted at the winder into 5 rolls. Thickness measurements and optical control systems for in-line quality control are integrated for quality monitoring. Four PLCs (S7), located at the dosing, the extruder, the nozzle, and the winder stages, regulate the 200 machine attributes, such as temperature and speed.

Data Acquisition

The 200 acquired dataset attributes from 4 data collectors and the sensor locations of laboratory machine M150 (Fig. 74) described in Chapter 3 (i.e. 140 million datasets since start of the investigation) are continuously transferred within 1 min by query connection (fieldbus, TCP/IP) to an Oracle database (MDE Server System). The machine is linked to a locally closed local area network within a laboratory hall, which, for prototype purposes, is separated from the productive systems.

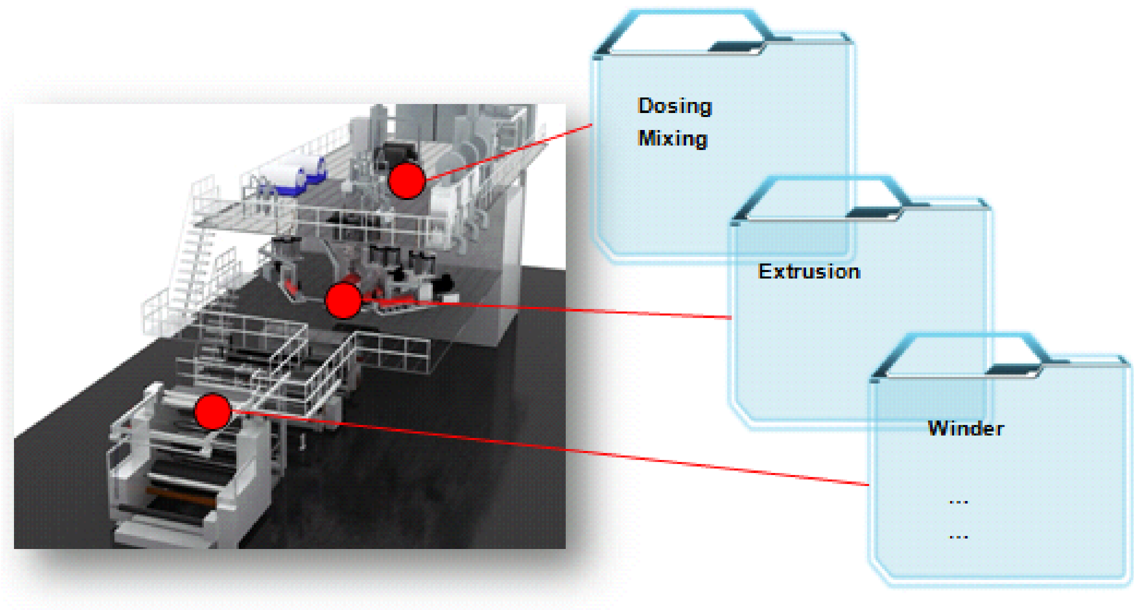


Figure 74 Laboratory *Cast*-extrusion machine for experimental monitoring investigation

A second query from a laboratory computer (Windows 7, 64 bit, Intel i7, 3.40 GHz, RAM: 8 GB memory, DDR3 SDRAM – 1600 MHz, Cache: 4 MB) collects a defined set of 160 mainly process quality contributing attributes from the MDE server system each minute for processing, amounting to 104 million datasets per year of several parameters such as speed, output, temperature, pressure, dosing, winder, optical defects, color and thickness values, for investigation.

Product Quality Measurement

Since the installation in September 2013 until now, about 4000 rolls were produced and monitored by an objective quality optical control system (OCS) connected to a manufacturing execution system (MES), providing statistical information about the production results (Chapter 3). Each roll is monitored by 2–4 (depending on the roll width) cameras transferring surface light gradient changes to a local OCS computer. A sample view is shown in Fig. 75.

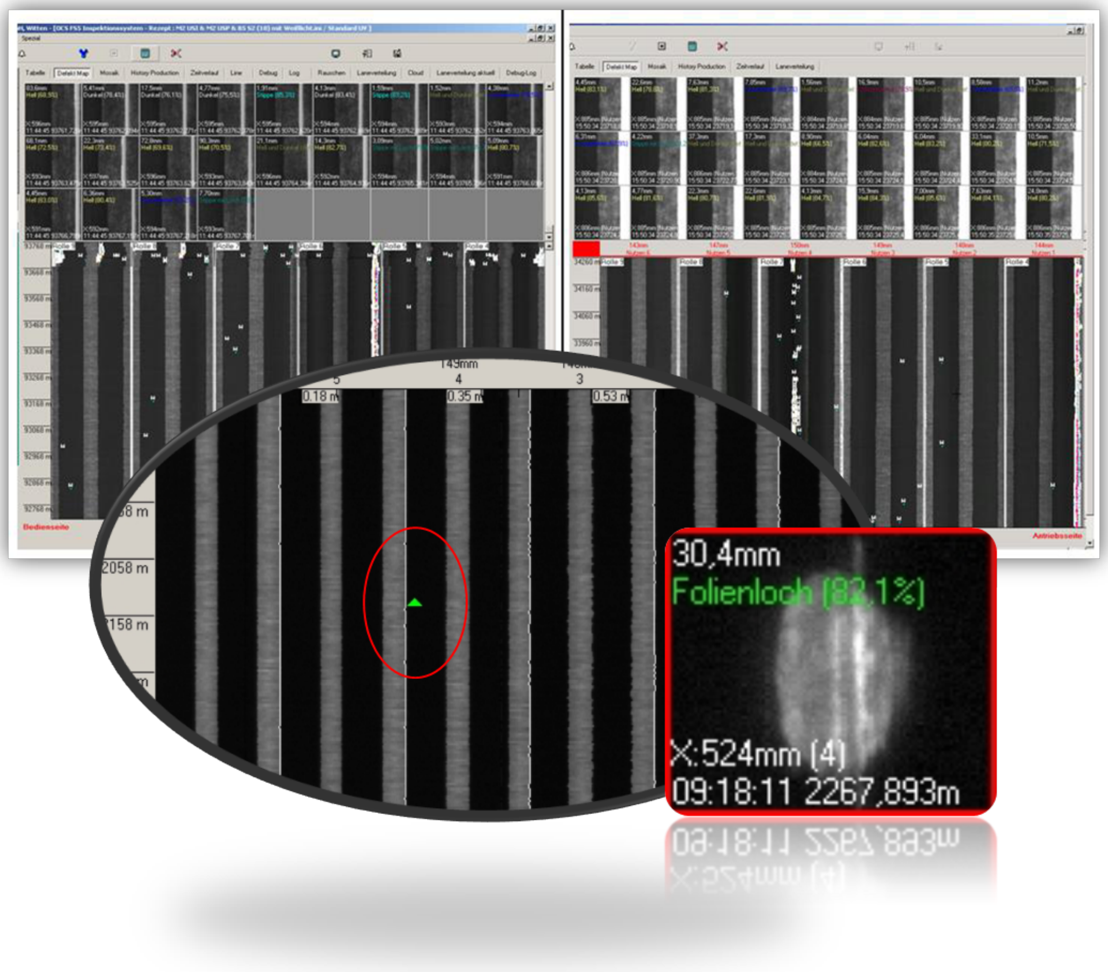


Figure 75 Sample picture of the optical control system from the production machine, monitoring a complete process roll for defects as holes, gels, or contaminations

The optical detection system showed correlating trends compared to the novelty and trajectory monitoring. Uncritical defects increased slightly evident as changing machine attribute values that showed up before the occurrence of quality deterioration and waste. The OCS software detects different 2-dimensional defects by geometrical investigation and classifies them by k-nearest-neighbor or decision tree into one class, as e.g., gels, holes, holes + gels, or contamination. The classes are trained by the staff with an uncertainty of 10% due to subjective allocation, with a typical defect classification accuracy of more than 80%, which is the minimum accuracy according to customer specifications. The usual defect classification archieves 90%.

A calculated quality ratio (QKZ[33]) from the OCS is transferred to the MES system (Chapter 3). This way, normal and abnormal condition states for each roll are objectively stored for OCC post mortem analysis by the quality management department. Beside the OCS, similar systems are available on the market from, e.g., Isra Vision, Cognex, or Keyence.

By transfer of the QKZ from the OCS to the MES, the traceability of the products is guaranteed, and the staff is supported by statistical quality trends of product types, as shown in Fig. 76.

[33] Classification ratio, developed by M. Kohlert to give quality information on rolls moving from one stage to another

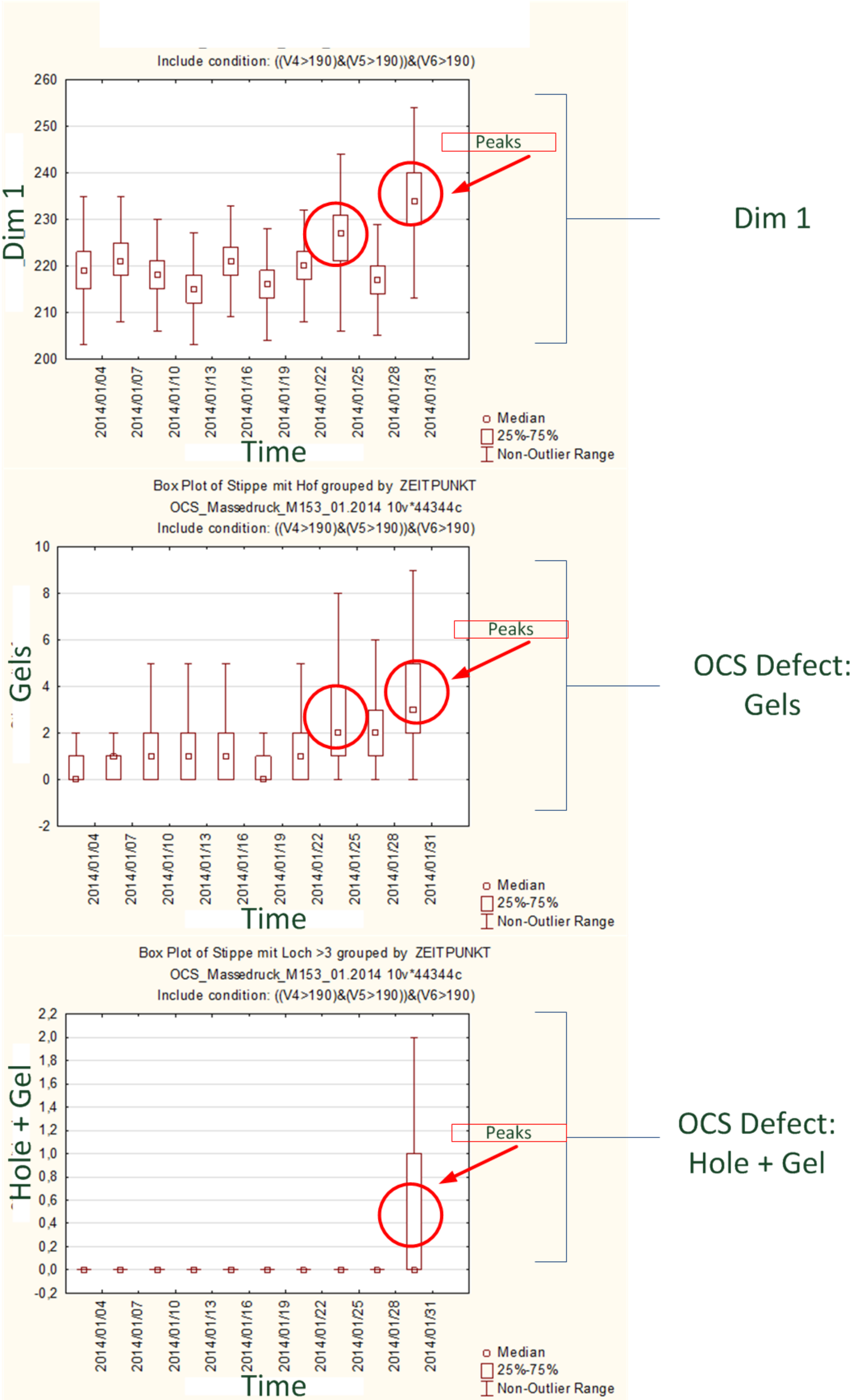


Figure 76 The optical and process data analysis from January 2014, showing uncritical correlations between pressure problems and optical defects on the finished roll

In Figure 76 typical processing attributes are displayed over time. A direct correlation between the Dim 1 (subplot top in Fig. 76) and the OCS defects (gels and holes, subplots middle and bottom in Fig. 76) is presented, by comparing the boxplots. Such relations are necessary to come to a second conclusion from an objective monitoring optical control system, in order to compare the predicted incidents with the final result. Thus, the optical defects have low impact, providing no direct indication of the specific "cut-off" problem.

Furthermore, the optical defect detection step is located at the end of the machine at the winder, and process problems are monitored belatedly post mortem at the process end.

In contrast, the "pressure" is directly located in the middle of the extrusion process. By observing the pressure, the main quality incident, "cut-off", which cannot be recognized by the OCS, was predicted.

Novel Monitoring Environment

According to the approaches detailed in Chapter 7 and the operator interview, the methods were integrated into a graphical user interface on the laboratory machine, to permanently visualize the results at the machine terminal computer for direct monitoring, which has not been established before. Therefore, the ideas from the operators' feedback were applied. They are described in the following paragraphs with regard to an extension to the productive online machines as outlined in Chapter 9.

The realization within a process monitoring environment was accomplished according to the system architecture described in Chapter 6. The layout for the software, constructed as shown in Fig. 77, includes the previously described settings, such as the data acquisition part, the classification methods, the software interface, the processing computer hardware, the visualization elements, and the message system.

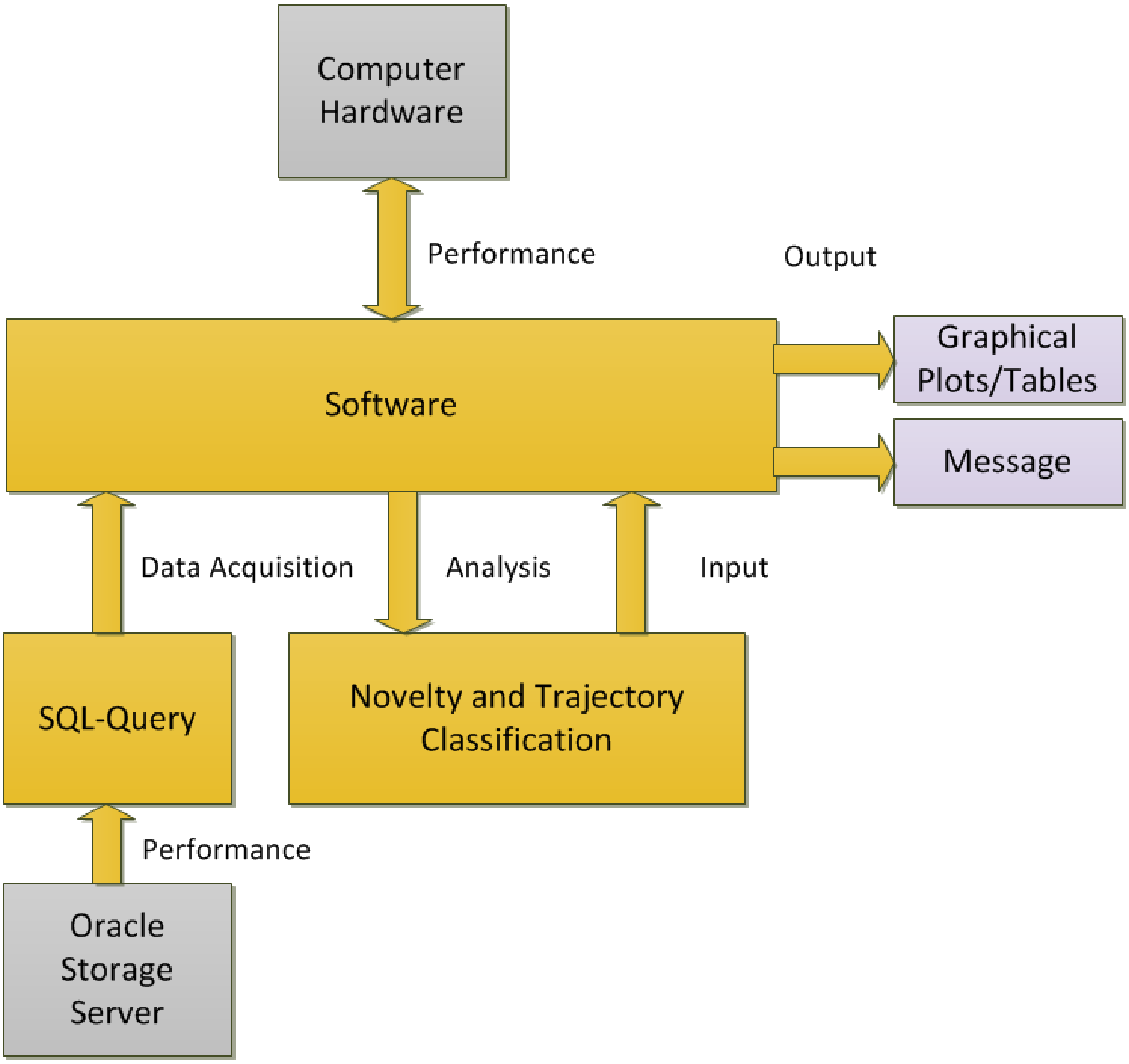


Figure 77 Software diagram overview according to the system architecture from Chapter 6

The data acquisition and processing parts (Chapter 7) strongly depend on the performance ratio of the storage system and the computer hardware used. In-Memory storage technology and a 64-bit computer system are recommended to achieve fast acquisition and recognition results, visualized and transferred within 1 min for at least 200 attributes such as pressure, temperature, speed, output, order number, recipe, and winder settings, and calculated attributes such as pressure gradient and amplitude, which are mostly double and string dataset types (see Chapter 2).

8.2 Realization within novel Process Monitoring Development

The whole real-time monitoring cycle is displayed in Fig. 78. It consists of the machine sensor locations, the PLC, the data collector, the MDE server, and the on-line laboratory computer interface connected by fieldbus and TCP/IP (Data Stream).

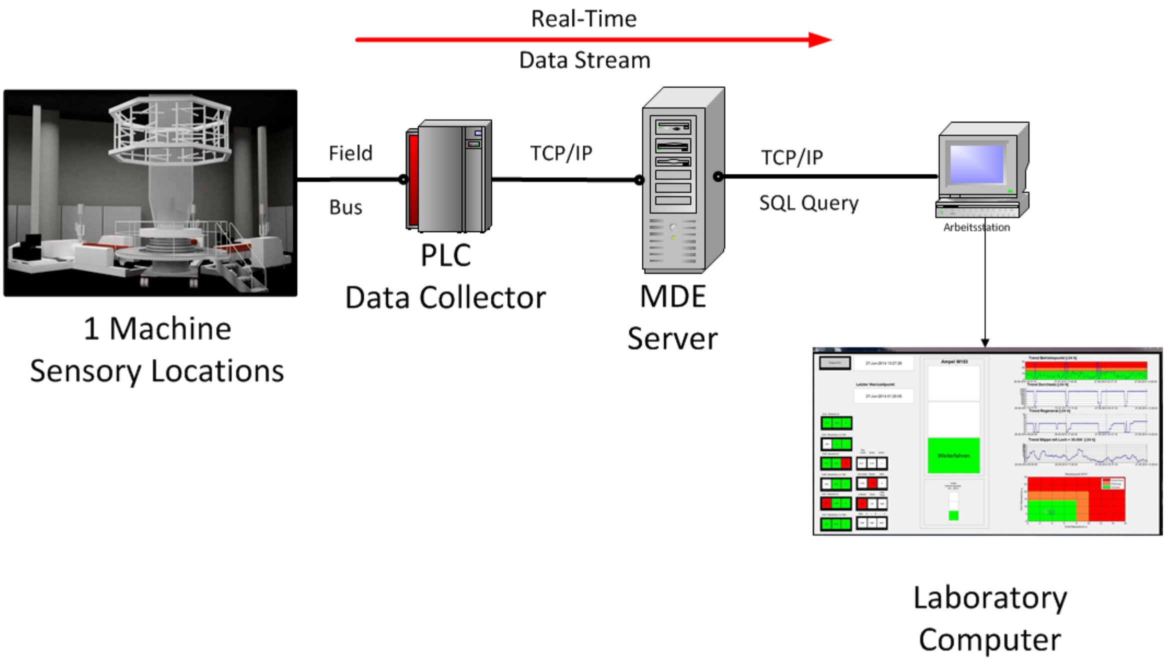


Figure 78 Laboratory data acquisition condition real-time monitoring system: machine sensory locations, data collection via plc, transfer to MDE server storage via field bus, monitored at laboratory computer

The trained OCC method classifies the new life datasets in 6-dimensional space and monitors the characteristic elements for further trajectory classification, as class and accuracy, in the laboratory interface programmed with *Matlab* & *Simulink*.
The prototype application gives recommendations on the different states of the machine by light and message box, as presented in Chapter 7.
The proprietary application-specific graphical user interface, developed and compiled in *Matlab* & *Simulink*, as shown in Fig. 79, allows the laboratory staff to on-line monitor the attributes, classification measures, and limit violations from the laboratory machine. The current attribute values are displayed in 30 text/edit fields and 4 time series plots: 1 scatter plot, in which the latest warning date is given as well as the current and the two previous ones; the recognition points are displayed in a 2-dimensional scatter plot, which is the Sammon's Mapping scaled view previously described in Chapter 7. The Sammon's Mapping recall projects the scaled 2-dimensional data points onto the scaled mapping plot [Kö00]. In this scatter plot with colored traffic light zones, the staff can track the current condition changes, the trajectory is calculated. The aligned elements are fixed on the interface surface.

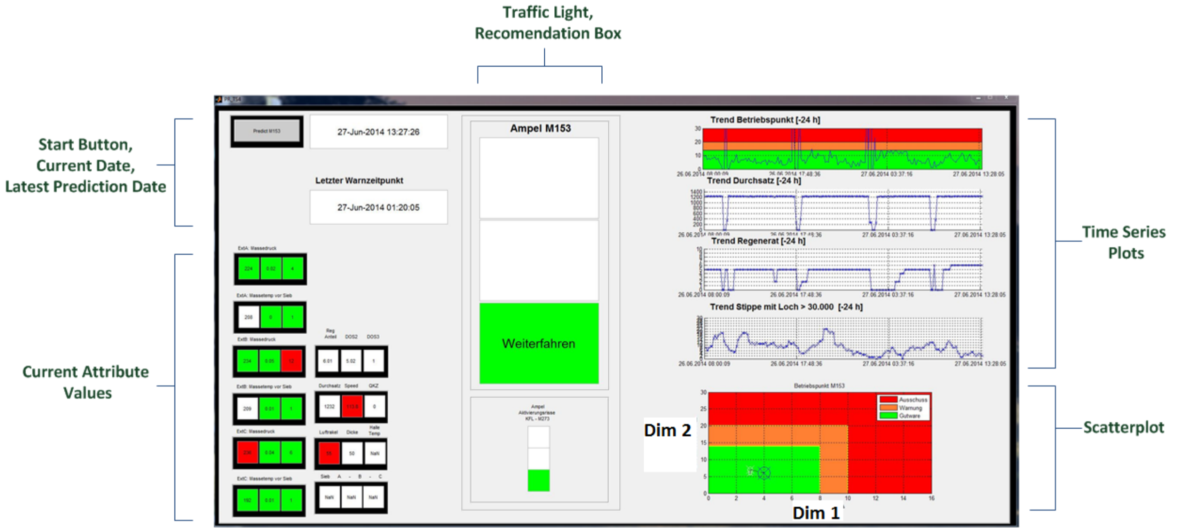


Figure 79 Graphical user interface for process condition visualization in the polymer film industry at the Mondi Gronau GmbH

The recommendation system is established with a e-mail system, sending messages about the current state within 13 s (for the laboratory prototype version) to the machine operator, the supervisor, and the division manager. Mailing is delayed after the first warning, to prevent message overflow.

The following practical incident is described by the current implementation in Fig. 80. The trajectory monitoring system (1) predicts an incident at about 60–120 min in advance, displayed by the upper plot as a green line. Beside the scatter plot in the lower-right corner showing “Dim 1” and “Dim 2”, each current scalar point can itself be displayed per minute in a time series plot (upper-right corner), as e.g., “Dim 2” per time, representing one dimension of the scaled data point from the lower-right scatter plot. This way, the operator can switch his attention from the previous scatter plot to the time series plot, to watch changes leading to final waste or the normal condition over longer periods, e.g., 24 h, as described in the previous chapter. In fact, the previously described Sammon’s Mapping from dim 6 to 1 instead of 2 has been achieved for visualization, which will be kept in mind for further methodological improvement.

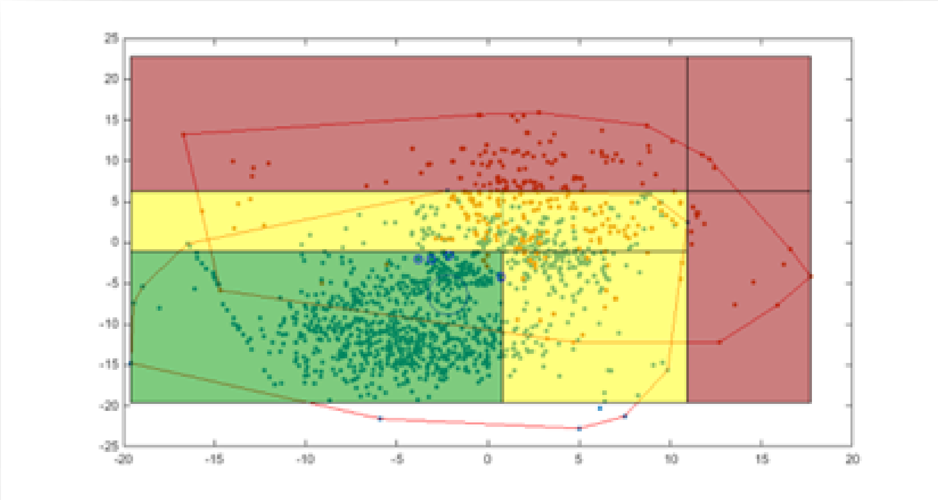


Figure 80 Sammon's mapped multi-dimensional processed 2-D Plot

The border lines are implemented to demonstrate the limit violation to the staff, according to the zone separation from the Sammon's mapped view after Delaunay and threshold setting (Fig. 58 in Chapter 7), which leads to a recommendation message regarding a possible "cut-off" problem. But the threshold does not indicate the real classification boundary between the two classes.
The "cut-off" problem includes all rolls that do not achieve the complete length of about 8,000 meters due to a big hole (a cut from material instability), dividing the roll during the production and stopping the process. A typical normal roll is produced in one piece without any big holes or cuts, and without a stop. If a "cut-off" occurs, the roll is divided into, e.g., 4,000 and 4,000 m. The stop and restart of the process takes a long time and is very expensive due to rework. The rework checks these roll parts for defects.
The implemented classification method calculates, in the 6-dimensional space in every minute, the class of the current trajectory for becoming a waste or a normal product roll, with the normal or abnormal class being visualized later by a traffic light system (Chapter 9, page 150). The OCC method is set as the main classification part, for as long as the trajectory has not been calculated. Each time the trajectory is calculated, the OCC result is replaced. In further approaches beside this work, the trajectory will be calculated at any time and the staff will be able to switch between classifications by OCC or by trajectory.
In this case (Fig. 80 16.02.2015: 03:44 am), a recommendation was sent to the staff and the machine was not stopped or adjusted by the staff. A "cut-off" (2) occurred, which by itself stopped the process, and the roll had to be reworked on the rework stage. As investigated in Chapter 7, the incidents were only monitored; no actions were taken.
For recommendation purposes, the staff receives a message from the system to take actions, such as executing a controlled "shut-down" of the whole machine. The machine speed is then decreased and the output is down-scaled in cascades, combined with

cleaning tasks. Thereby, the “cut-off” and approximately about 60 min of additional cleaning time are avoided.
In a first selection, the NOVCLASS method by König and Gratz (2005), the neural networks auto-encoder, and the *k*-nearest-neighbor classifier corresponding to Tax (2013) were chosen for integration into the Matlab & Simulink source code, due to their simple modification options.
The graphical user interface for the visualization of 5 time series plots, 1 traffic light, and 30 current value fields was generated as a prototype. The incident is represented by a traffic signal in red, a message box recommends the best adjustment on the screen, and an e-mail is sent to the registered interface user.
In the following, the implemented summarized complex functions and their GUI locations are listed:

- Traffic Light (Middle): OCC Classifier
- Scatterplot (Lower-right): Sammon’s Mapping & Recall
- Value Box (Lower-left): Trajectory Classifier
- Message Box (Middle): Message System

8.3 Prototype Assessment

The assessment efficiency described by measurement ratios from the plant (e.g., material waste, critical roll count, or savings) need to be calculated objectively by a process observing system. The manufacturing execution system (MES) stores the quality and processing efficiency data for all machines (also the laboratory machines), to reproduce ratios for downstream analysis.
In the MES, each produced roll acquires a marker indicating at which time a specific problem has occurred. The marker is correlated to a quality ratio called QKZ as explained above, giving a recommendation via display to rework, to sell or to destroy the roll after the roll production process is finished. In the next step, the staff transports the roll to the specific locations at the plant, e.g., the rework, storage, or destruction area. Each step is deposited in the MES as digital numbers and description.
Afterwards, at the end of the year, the Controlling department analyzes the complete year by automated calculations, collecting all accumulated datasets of all rolls at all machines. This way, at the end of the year, an annual financial statement is made and every division and machine is benchmarked regarding waste (%, kg, EUR) and performance compared to the previous year, as ratios.
Since the installation at the machine terminals and the office computers, the waste ratio – calculated by the Controlling department (which is the main objective institution at the plant) – was reduced from about 20% to 12% at the prototype machine, as presented in Fig. 81. Although the material input increased in October 2014, the waste trend decreased by up to 8% and the variations were lowered; the time series marked in red indicates the waste percentage and the one marked in blue represents the number of rolls with “cut-off” problems. The mean waste amount calculated from the absolute waste amount per month decreased since installation of the system, with a difference of 2.72% when comparing 2013 with 2014, and 0.55% comparing 2012 with 2014.

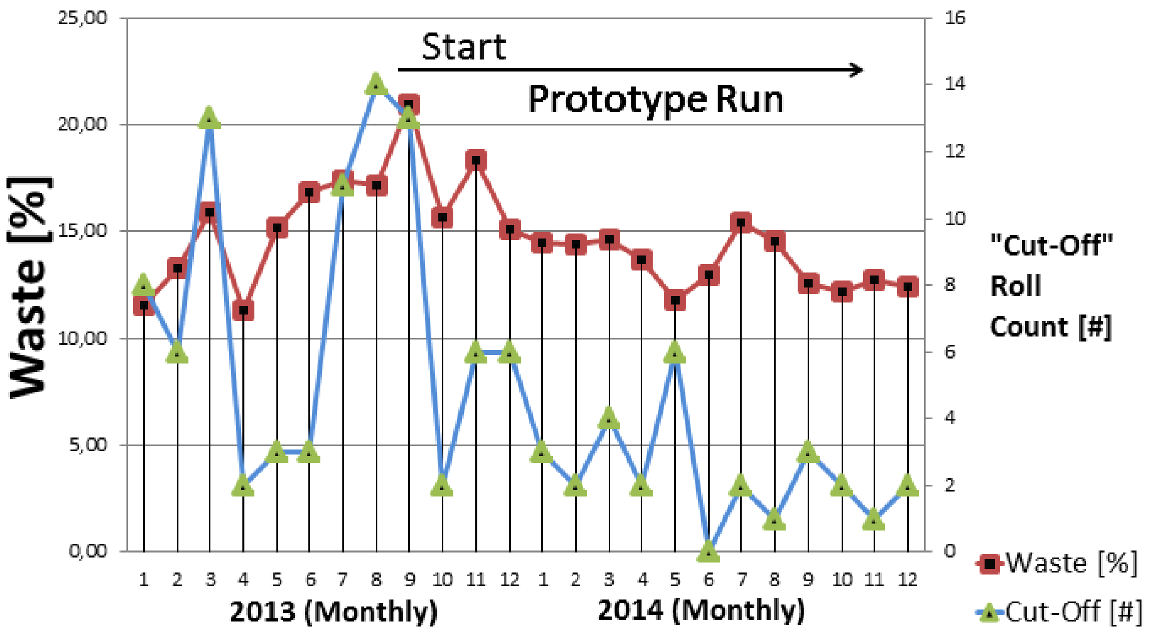


Figure 81 The decreasing material waste trend for the specified extrusion line M15X extracted by the Controlling data from 2013 to 2014, marked red, and the decreasing material critical number of rolls trend and amplitude extracted by process experts from MES

The decreasing material waste trend for the specified extrusion line M15X extracted by the Controlling department from the data of 2013 to 2014 data (marked red) and the decreasing trend in the critical number of rolls and amplitude extracted by process experts from the MES are displayed in Fig. 81. The prototype started the classification

Table 25 Summary of Results

Compare Type	Unit	2012 vs. 2014	2013 vs. 2014	Sep_2013 vs. Sep_2014
		[Mean]	[Mean]	[Absolute]
Waste Material	[%]	-0.55	-2.72	-8
Waste Material	[kg]	-1,380	-6,800	-20,000
Waste Rolls	[%]	-2	-10	-30
Waste Rolls	[#]	-3.8	-18	-55
Savings	[%]	-0.03	-0.17	-0.5
Savings	[EUR]	-3,450	-17,000	-50,000

and recommendation in September 2013. The analyzed data was taken from 2013 and 2014, for a better comparison. In addition, the staff spot-tested the prototype with a recognition rate of 90% of correctly classified incidents in advance of occurrence, counted by the machine operator from September 2013 to September 2014.
The implementation was chosen from the main elastomer division due to the 70% yield share of the division's operating results. The final yield evaluation from the Controlling department decides on further production at the machine or a product change.
A teaching time of 2–4 weeks was necessary to explain the monitoring & recommendation system to the staff performing the manual adjustments at the processing machine. Therefore, from the beginning in September 2013, the decreasing trend set in with a slight delay, in October 2013. Simultaneously, the variations were reduced and, in sum,

the process switch from OCC to trajectory classification is changed now.
The manufacturing execution system stated a critical roll count reduction by 30% for the specific problem, which corresponds to an absolute count of 55 rolls. The improvement of the specific “cut-off” waste problem was achieved during the period from 2013 to 2014; this improvement could not be anticipated before this investigation .
The critical number of rolls decreased during the observation period, although a correlation to the previous Controlling evaluation is not possible due to different viewpoints. In this part, the number of rolls with the specific problem is regarded, which, in contrast, means that the waste amount does not decrease that much, due to rework and subsequent waste problems on the following rework machine. The system detects the specific “cut-off” problem in advance to avoid further processing time and reduces the number of critical rolls per month by bringing the machine to a controlled shut-down. A uncontrolled random shut-down leads to readjustments of the machine settings and a cleaning phase combined with at least 30 minutes of time effort, which could have been avoided by earlier condition classification. The affected roll then needs to be reworked on a slitting machine to cut out the critical parts, e.g., about 20% of, e.g., 4,000 m. Such short rolls are rejected by the customers; therefore, selling them is difficult and leads to yield losses of about 40%.
During the prototype testing and implementation phase *savings* of about 20,000 kg of material (= absolute) and EUR 50,000.00 corresponding to 0.5% of the total (including rework) waste cost, and an average speed increase by 2,000 m/h were achieved, measured by the MES and the Controlling department. Beside the absolute results, the mean values correspond from revised extrapolations. A first approach for a monitoring and recommendation system achieved yield optimization and can furthermore be extended to more machines, thus offering 5.0% more potential for material waste savings, gains of about EUR 800,000.00 by decreasing the total waste including rework efforts for the specific “cut-off” problem.
The investigated ”cut-off” problem made about 90% of the critical conditions. The remaining 10% of quality waste incidents, as e.g., color deviations, roll-up, or blow-up problems, will be investigated in future works beside machine energy savings.
The achieved results were based on real manufacturing execution and quality measurement systems. Computer systems, which are directly integrated into production for 20 years (the MES and the OCS system; Subsection 3.2.2), observing every produced roll at each second. The MES (manufacturing execution system) is the main objective plant production data collection system for all machines; it automatically stores the quality results from each roll, coordinates assignments of rolls to other machines and is objectively monitored at every second by the OCS (optical control system) located at the winder of each machine (see Chapters 2 and 3). Both systems are linked, automatically transferring information (datasets) about quality incidents and material flows. The quality change incidents are imminently correlated with the recognition results, which beyond that are verified through spot analysis by the responsible staff.
The presented verification of the on-line monitoring & recommendation loop with the implemented *Trajectory Visualization* was the decisive cue for projecting the examined approach to other extrusion lines directly connected to the productive system, for integration into the real-time production process.
Furthermore, the application-specific graphical user interface should be extended to more terminal interfaces, to give precise information on deviating process attributes of the complete division.
Regarding the whole plant, an estimated extrapolation of 80 tons of material waste

reduction per year could be achievable, transfered to the extensive polymer industry more than 8,000 tons per year could be conceivable.

9 Implemented On-Line Open-Loop Monitoring System Approach

Now, as the next evolutionary step, the investigated methods and approaches from the previous investigations, Chapter 7, including *OCC*, *Sammon's Mapping*, and *Trajectory-Visualization*, are integrated in the continuous production process.
The on-line prototype for real-time monitoring was installed on eight selected extrusion machines directly integrated into the production process and investigated with regard to the currently implemented system architecture described in Chapter 6.
The novel application-specific graphical user interface from the previous laboratory approach for near-real-time acquisition, processing, visualization, and recommendation, the so-called *On-line* ***Con****dition Real-Time* ***Mon****itoring System* (**ConMon**), was transferred in a preliminary step for testing and adaptation on eight *cast* extrusion machines (see Chapter 3), to provide added value for the running production.
The complete setup containing the processing hardware parts with **research & development production sw supplementation** were adapted to the requirements of the current products. They were directly included in the production process and connected to the internal machine network (VLAN). Due to their connection to an Oracle database, the so-called research prototypes stored datasets in 1-min intervals from altogether 16 machines at the plant. The preferred data stream cycle is displayed in Fig. 82.

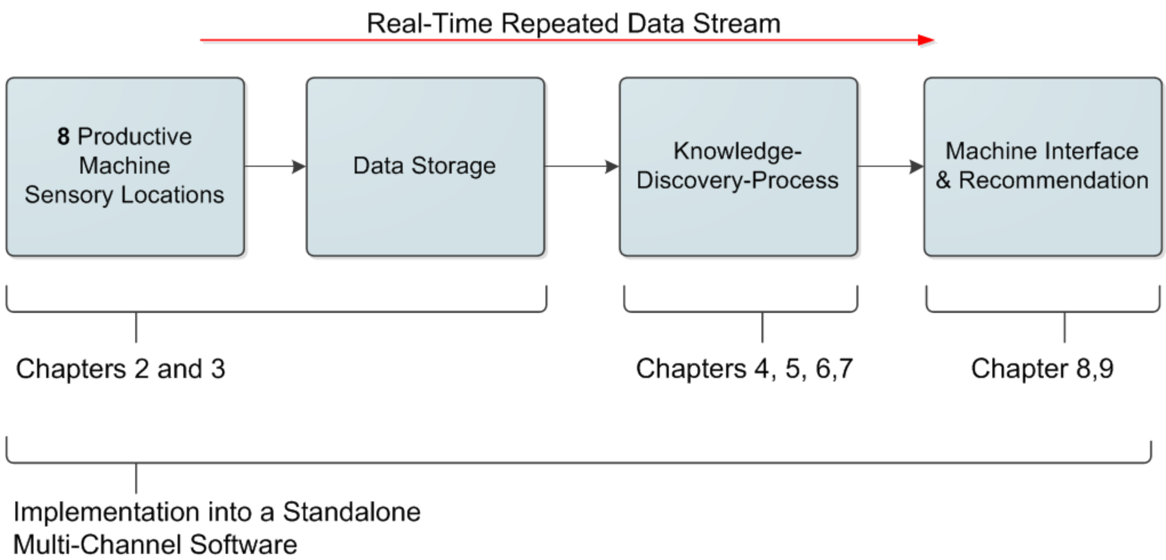


Figure 82 On-Line data acquisition condition real-time monitoring system: machine sensory locations, data collection via plc, transfer to MDE server storage via field bus, monitored by a standalone multi-channel software

The complete "Data Stream" from the laboratory machine sensor locations to the interface visualization is divided into three groups. The **first** group includes the sensor locations, the communication links, the data acquisition, the data types, and the storage settings, as described in Chapters 2 and 3. The **second** group deals with the dataset processing, feature reduction, and classification (see Chapters 4–5) based on the *Knowledge Discovery Process* described in Chapter 2. The **third** and final group regards the datasets and methods used and the result visualization and recommendations as requested by the staff (see Chapter 7). The complete setup was implemented into a real-time monitoring data acquisition & recommendation standalone multi-channel

software at the machine terminal and in the staff office.
The following sections are arranged as follows: (9.1) *Data Acquisition, Processing and Investigation*, (9.2) *Interactive Visualization and Recommendation System*, and (9.3) *Prototype Assessment*.

9.1 Data Acquisition, Processing and Investigation

The real-time monitoring of trajectories and conditions was implemented as a standalone application close to the manufacturing process. A monitoring interface at the machine terminal supported the operator with graphical evaluations and recommendations for process intervention. The process changes made were reported by business objects for subsequent downstream analysis. The complete architecture was established as a prototype version in September 2013; since July 2014, it has been implemented as an integral part of the plant monitoring equipment.
The development from the laboratory setup to the real-time online monitoring system for the production area was achieved in a second step. Eight principal cast extrusion lines for SBS, PE, and PP[34] (Fig. 83) were appropriately equipped to recognize patterns in the process conditions before the occurrence of abnormal states.

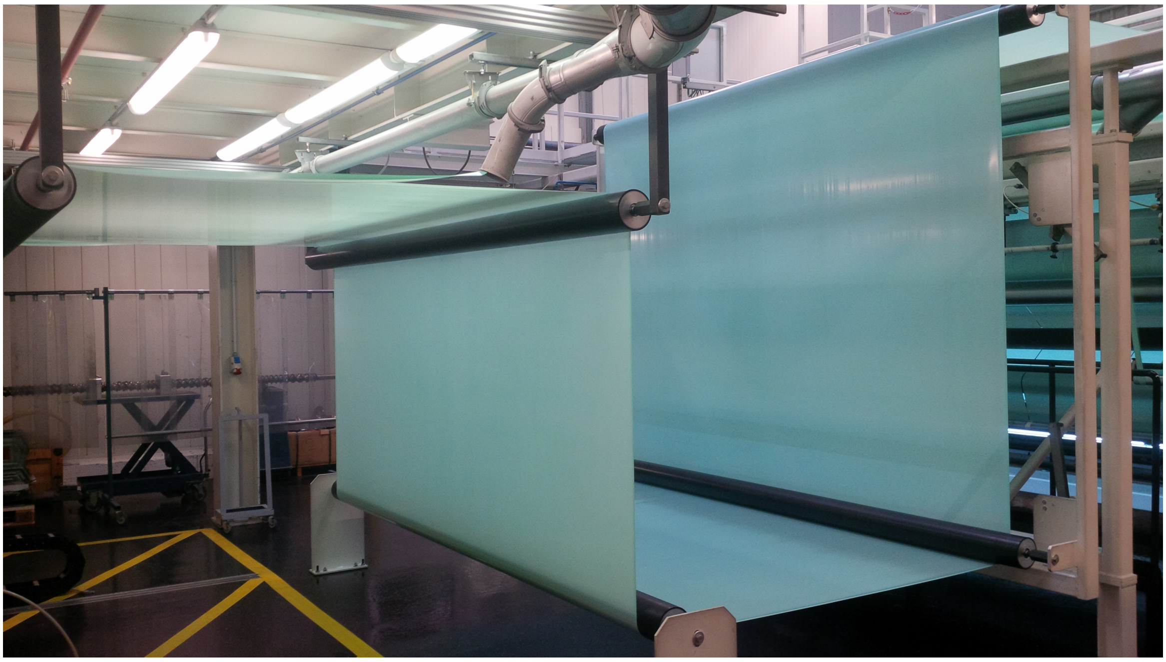


Figure 83 Cast-extrusion polymer process machine for the production of multilayer films

A typical cast extrusion machine produces films of different width and thickness for hygienic components, which are externally processed by the customer, i.e., the customer's plant.
The complete laboratory setting was transferred to eight terminal computers directly at the machine sites[35]. Besides monitoring, the terminal interfaces are mostly used for MES inputs, specification modifications, or communication. All terminals are directly

[34] Index of Abbreviations
[35] Additionally, driver and sql statements were upgraded for local and Citrix server application

linked to a machine network, which operates in parallel to the business network (Fig. 84).

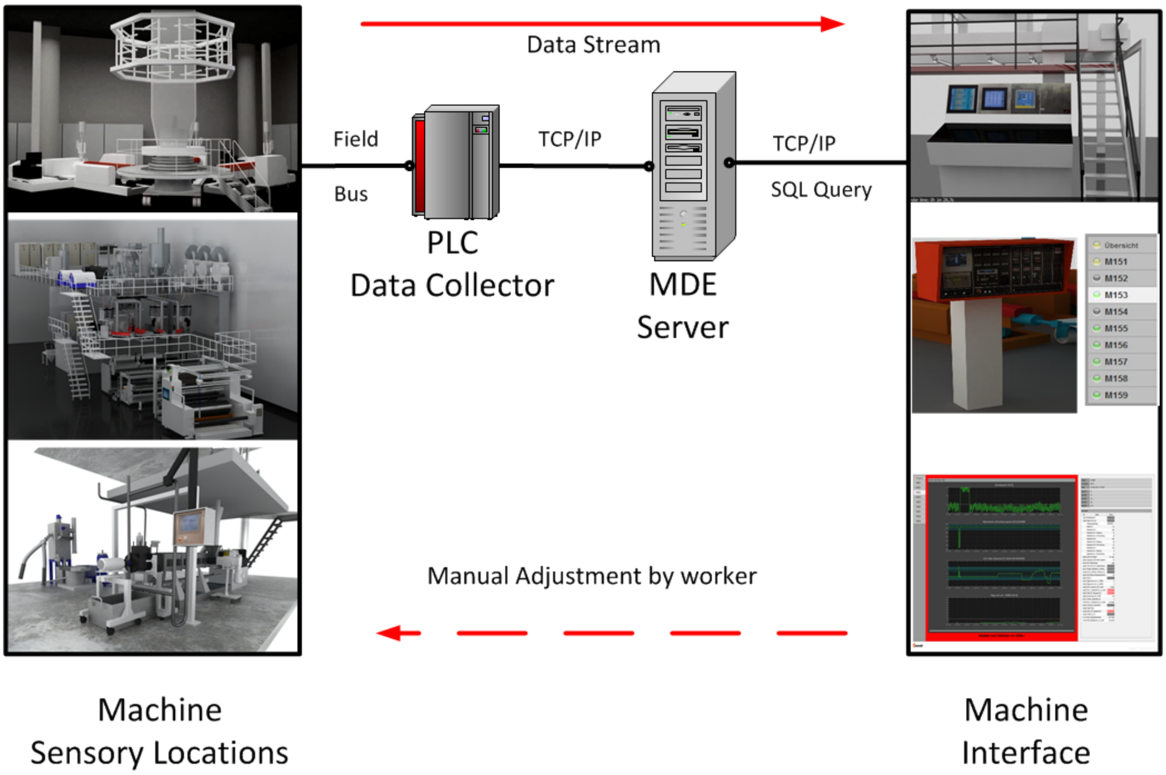


Figure 84 Online open-loop monitoring system from machine sensory locations to machine interface, realized at the Mondi Gronau GmbH

Both networks are connected and conducted in redundant mode for safety reasons with regard to malfunctions. The machine setup interface is not linked to the terminals. Newer machines integrate most of the setting options at the setup interface for a complete overview. Data acquisition is established via a PLC data collector, which each minute transfers the data to the MDE server (Oracle). The standalone application queries the datasets via TCP/IP, to be displayed on a machine interface for recommendation.
Depending on the recommendation type (described in the following subsections), the operator is always able to choose between two types of action: (1) doing nothing, (2) adjusting the PLC settings. In case of warnings, the staff starts with the second type of action by adjusting the temperatures of the main extruders (e.g., lowering by 10°C), decreasing the speed by 20 m/min, and monitoring the upcoming number of defects like gels or small holes at the optical control system. If the warning persists for about 15 min or if the critical condition is reached, the machine is stopped by lowering the speed to 0 m/min; the temperature decreases, and all other states, such as dosing, are stopped. Then, the machine is cleaned by opening the extruder and afterwards restarted until the previous production output is reached again. Altogether, this amounts to 30–60 min of stopping the machine(s). Under normal conditions, typical extrusion lines reach production times of 2 weeks without stopping. All subsequent machine modifications ("Manual Adjustment") are executed by the operator. Each minute, new datasets for all specified machines are generated and possible attributes are monitored (7 days back) through an Ethernet connection to the Oracle database. All methods from Chapter 4

were implemented and tested during the investigation.
In this presented case, the support vector data description method [Tax13] was integrated, although each machine contributes different attribute groups (from 40–500 sensor locations), due to their construction.

9.2 Interactive Visualization and Recommendation System

The standalone application was generated with Matlab & Simulink. The upgraded GUI for the observation of eight machines, as presented in Fig. 85, differs from the previous interface depicted in Chapter 8 with regard to the design. The functions are the same; thus, a drag & drop routine was additionally implemented, which allows the user to switch the presented data in the time series plots. The GUI offers an interactive user interface for real-time monitoring of the process datasets. The software shows details of the processing attributes for different time shifts (1–7 days), including statistical specifications and dynamical limits.

Figure 85 New graphical user interface for observation of 8 extrusion machines

A total view gives information on all machine condition states by traffic light and/or by sending recommendations via e-mail or SMS if the current dataset classification result changes from normal to abnormal conditions:

- Green: Good operation mode
- Yellow: Warning! Critical condition!
- Red: Shut down machine within next 3,200 m[36]

The operating condition is adjustable by changing the temperatures (170–200°C) in the extrusion zones, trimming lower settings for the speed or the torque, increasing the dosing, or decreasing the percentage of the returned re-granulate by the operator. At the machine terminals, the staff is able to monitor the current process data and receives recommendations in case the condition state changes in the next few minutes (Fig. 86).

[36] Grey = Offline

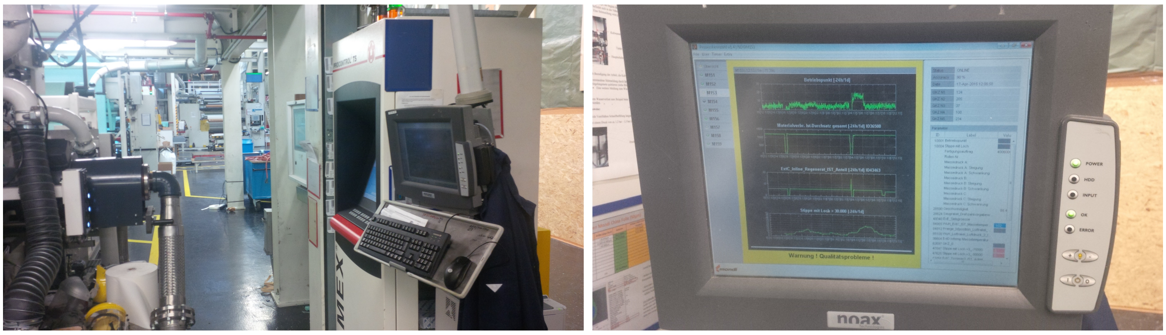


Figure 86 New graphical user interface located at the machine terminal M153

The same computer terminal located at the main machine observation area is also used for MES input. At this point, the staff adjusts the machine settings or receives information from the Planning department for upcoming processing orders. The machine types are visualized in the GUI by the traffic light system. This way, an overview of all machines at the same time is given at each machine, as displayed in Fig. 87. The possible condition states were mentioned above.

Figure 87 Traffic Light Visualization for Monitoring of eight Extrusion Lines, displaying their Condition States

The OCC and trajectory classification were extended to all machines and are calculated for all machines in each minute or later, depending on the refreshing rate set by the staff. If machine states change from green to yellow or red, a message system informs the user via e-mail and a message box appears, giving the previously described recommendations. The following summarized, complex, activated functions are shown:

- Traffic Light (Upper-Left): OCC Classifier
- Message Box (Middle): Message System

In addition, the following summarized, complex, not yet activated features are shown:

- Scatterplot (Middle): Sammon's Mapping & Recall
- Scatterplot (Middle): Trajectory Classifier

The new Drag & Drop function gives the staff the opportunity to interactively change the features displayed in the middle plots. As a novel add-on, the trajectories of all eight machines exist and are in use into the 2-dimensional Sammon's view, either separately or as a combination of all trajectories.

9.3 Prototype Assessment

The complete assessment loop is presented in the following as a prototype sequence. Therefore the polymer film process and the main location for the "cut-off" incident are regarded. The sequence includes recommendations for the operator to categorize the rolls by quality ratios to be explained in the following with a summary of waste reduction results.
New datasets from all machines pass the adapted interface at the same time each minute; for the year 2015, about 700 million datasets will be evaluated in sum, at the end of the year, for the specific waste problem. The typical "normal" polymer processing line at the main location where the "cut-off" incident may occur is shown in Fig. 88. At that point the film splits into two parts during abnormal condition states.

Figure 88 Polymer film process product roll main location for "cut-off" incident after extrusion and before roll-up

The performance ratio for each machine and division will be measured at the end of the year by the Controlling department, after collecting the quality results of each roll in each month from the manufacturing execution system, as described in Chapter 8.

Each produced roll receives a marker in the MES, indicating at which time a specific problem has occurred. Typical rolls from normal production are shown in Fig. 89.

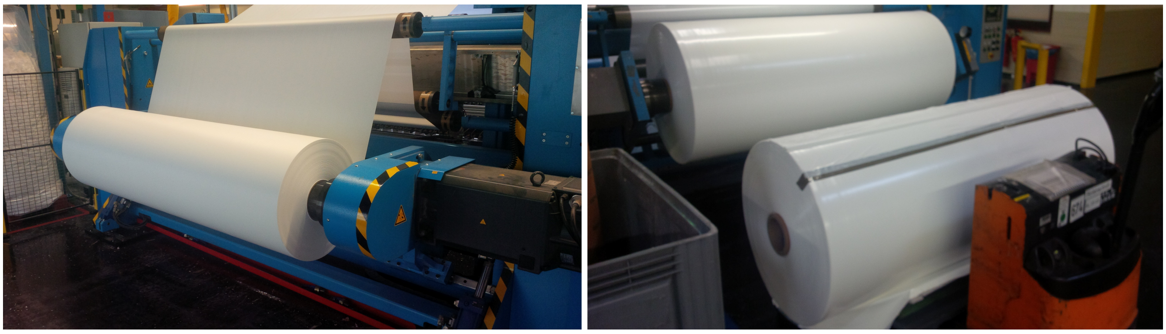


Figure 89 Rolls from polymer production process categorized as "normal" quality ratio

The marker is correlated to a quality ratio called "QKZ" as explained above, giving the recommendation to rework, to sell or to destroy the roll via display after the roll production process is finished. A roll with abnormal behavior during production ("cut-off") is displayed in Fig. 90.

Figure 90 Rolls from polymer production process categorized as "abnormal" quality ratio

In the next step, the staff comprising about six operators rotating between the eight machines transports the roll to the specific locations at the plant, e.g., the rework, storage, or destruction areas. Each step is deposited in the MES as digital numbers and description. In Fig. 91 the MES batch tracing interface offers information about each roll and material type from all machines with quality ratios.

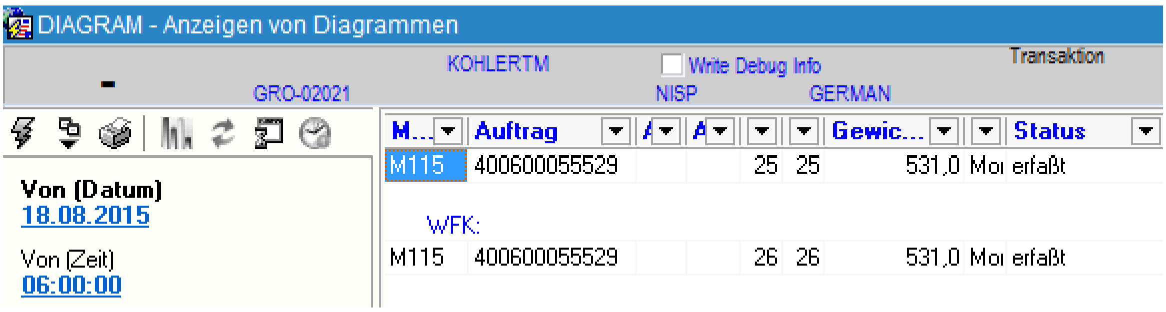


Figure 91 Exemplary MES batch tracing interface with material orders per machine and the quality ratio defined as "Status"

The quality ratio (the status attribute "erfaßt" means "normal" in this case) is stored within the MES database (Oracle) for further investigations. Figure 92 shows the locations of abnormal and normal roll products in the Sammon's 2-dimensional Mapping according to the quality ratios in MES for the specified rolls.

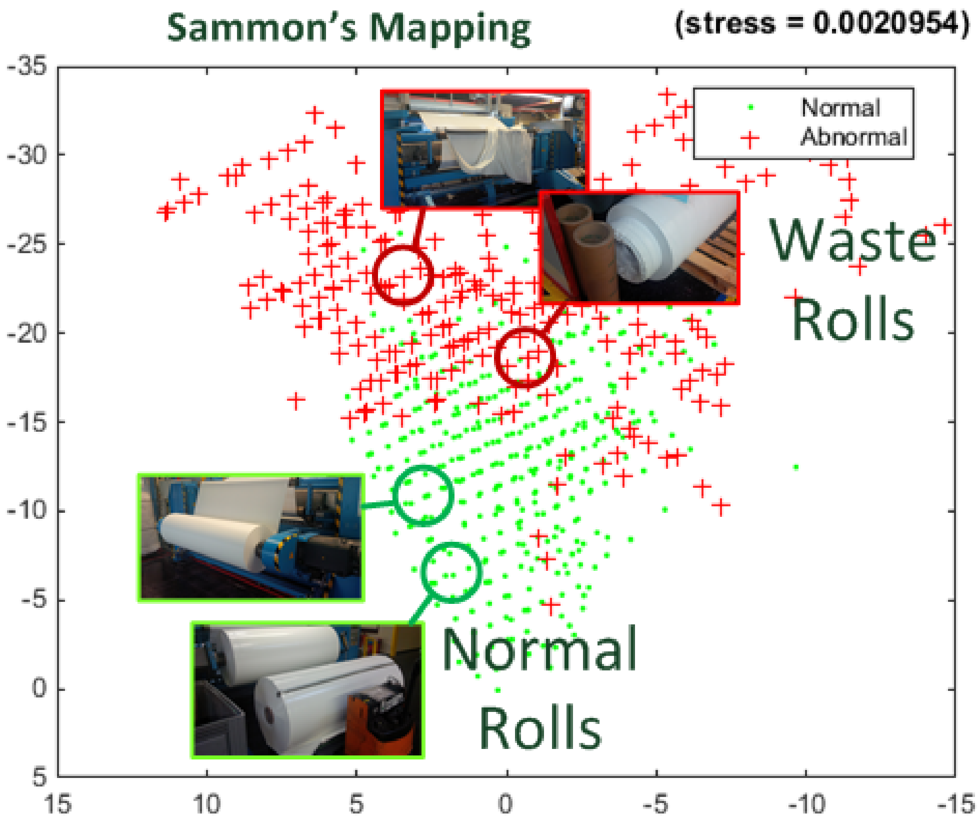


Figure 92 Roll product relation to Sammon's Mapping 2-D visualization

Each point of the 2-dimensional Sammon's Mapping plot describes the visualized condition of the roll with its quality ratio "normal" or "waste", as stored in the MES. The following figures give an impression of the four different possible sequences, starting with the first **(1)** "normal" condition (Fig. 93):

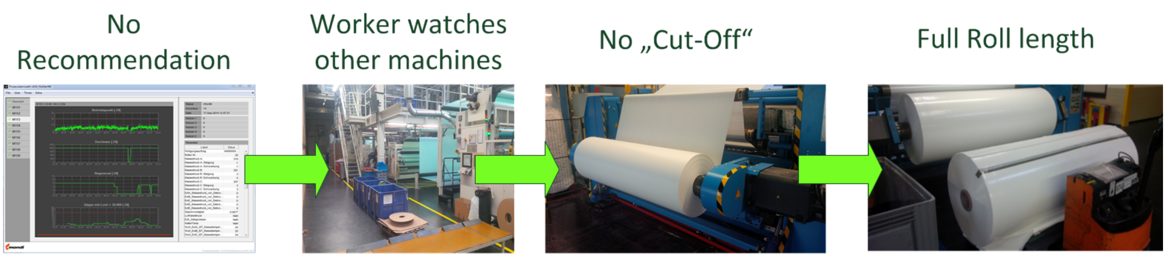


Figure 93 Normal sequence: no recommendation, no "cut-off", full roll length produced

The software offers no recommendation; therefore, the operator watches other machines and the production line generates rolls of full length.
Fig. 94 **(2)** shows the sequence in case of a warning condition, visualized as a message

box on the interface and as an e-mail on the operator's cell phone.
The warning occurs, when the current timestamp enters the boundary zone of the 2-D Sammon's Mapping. We remember Chapter 7, the boundary zone is part of the "normal" zone but with trends to the "abnormal" area. A warning is displayed by entering, but the condition is still "normal", the process is running at normal speed (120 m/min), an attribute which is monitored by MES for declaring a roll process as "active". The roll production is declared as still good, when the process is still running since start without interruption, checked by the staff.

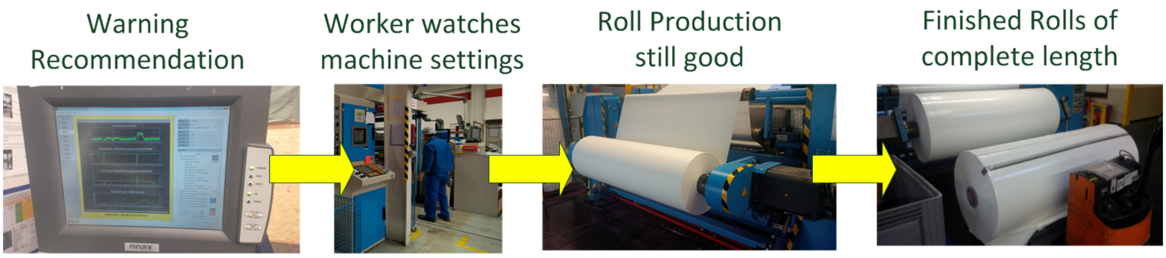


Figure 94 Warning sequence: warning recommendation, no "cut-off", full roll length produced

Fig. 95 **(3)** shows the sequence in case of a waste/stop condition, visualized as a message box on the interface and as an e-mail on the operator's cell phone. The process is declared as "inactive".

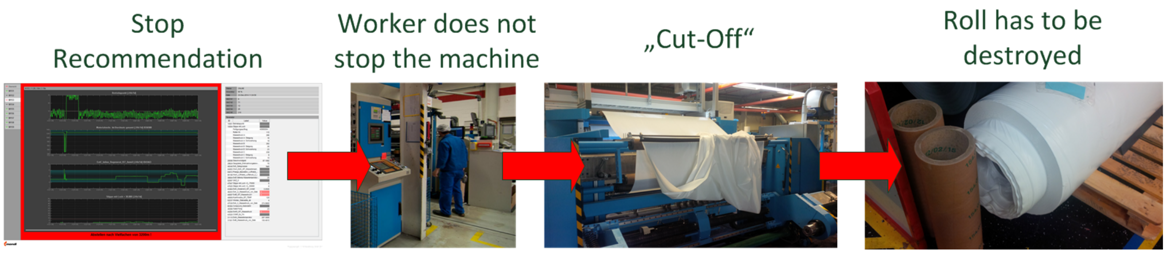


Figure 95 Stop sequence: stop recommendation, operator does not stop the machine, "cut-off", roll has to be destroyed

Fig. 96 **(4)** shows the sequence in case of a waste/stop condition, visualized as a message box on the interface and as an e-mail on the operator's cell phone.

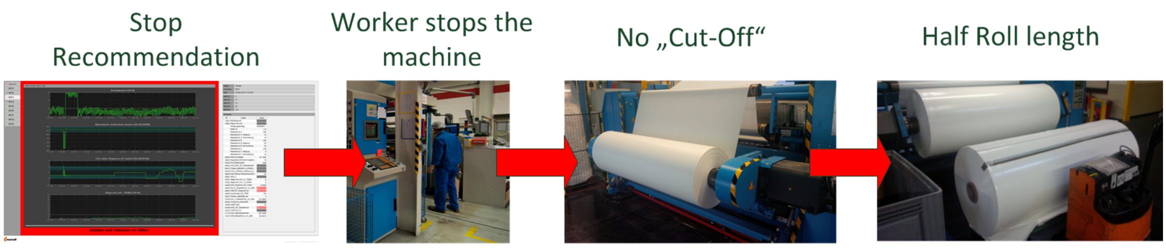


Figure 96 Stop sequence: stop recommendation, operator stops the machine, no "cut-off", half roll length directly sold to customer or sent to rework

The above sequences showed the different types of process development for "normal" and "abnormal" conditions. Fig. 97 combines the real data points from the Sammon's Mapping with the "normal" material processing conditions.

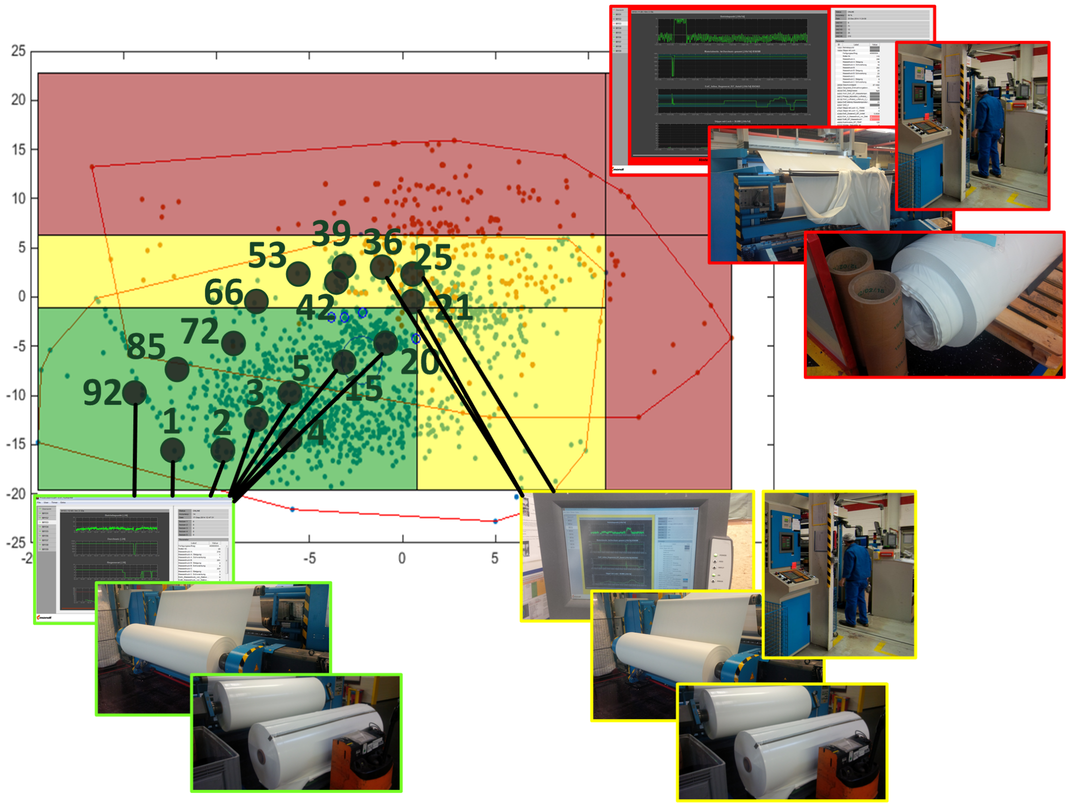


Figure 97 Normal processing cycle starting in good condition mode, running into the warning zone and leading to good production results

The timestamps are named, forming a cycle for one roll, 92 timestamps in the shown case, from the start to the final product without "cut-off". Fig. 98 shows the real data points of processing combined with the machine condition leading to material waste.

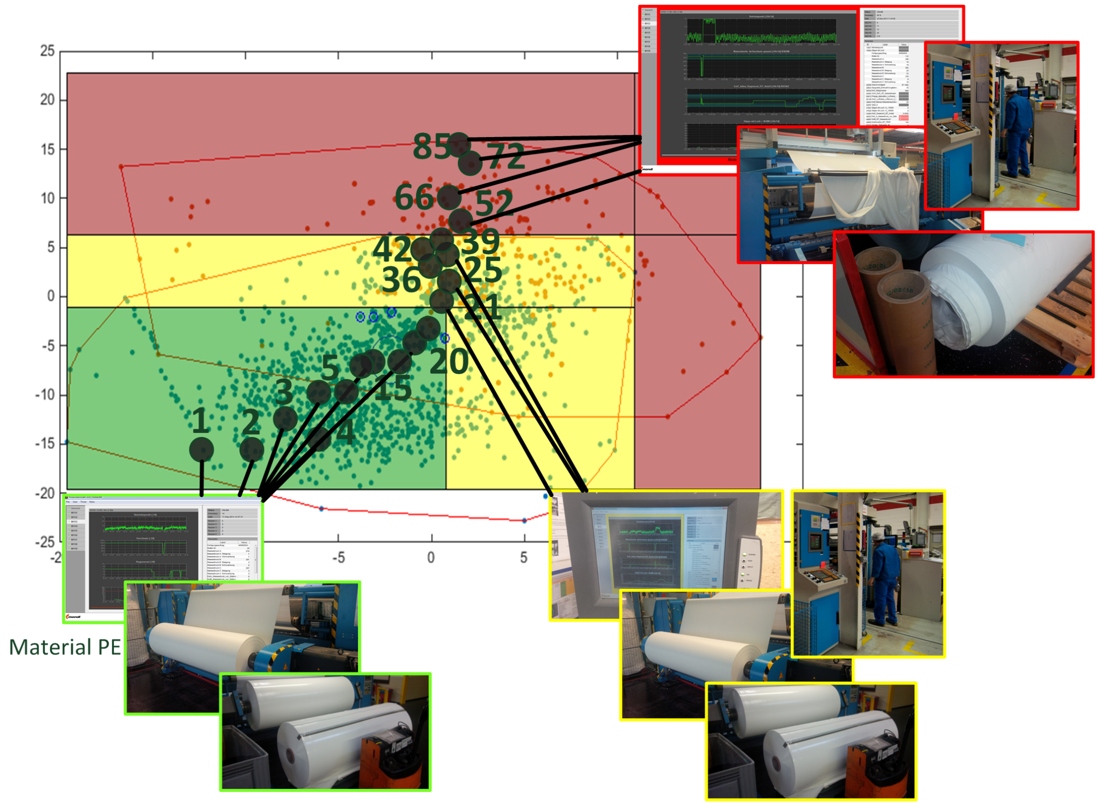


Figure 98 Abnormal processing cycle starting in normal mode, running into the warning zone, abnormal zone and leading to material waste due to ”cut-off”

At the end of the year, the Controlling department analyzes the complete data of the year by an automated calculation, collecting all accumulated datasets related to all rolls at all machines. This way, at the end of the year, the annual financial statement is made and every division and machine is benchmarked with regard to waste (%, kg, EUR) and performance compared to the previous year, expressed as ratios.
Such ratios need to be generated over several months to make sure that different types of products (product mix) are considered. A typical product mix cycle takes about 1 year per machine.
The waste and reduced number of “cut-off” rolls will be analyzed. Spot tests for correct or incorrect recommendations over 10 days per month will be executed by the staff to check for correctly recognized incidents in advance of occurrence, which in sum is at about 75% for all machines at the moment. The assessment is conducted with a simple check list, offering the following options: predicted incident *occurred* or *not occurred*.
Other influences on *not occurred* “cut-off” incidents are given by, e.g., unknown changes of the material properties caused by the supplier or by the transportation. Such options are not traceable due to missing or insufficient supplier processing datasets.
After these spot tests, the staff was ordered to permanently follow the instructions of the software recommendation system for another 10 days. Then, during the last 10 days of the month, the results were discussed. This way, a routine has been established since January 2015 and the software can be adapted repetitively for each machine.
An evaluation of the recommendation accuracy will be made at the end of the year in a holistic analysis by the Controlling department. Since installation at the machine terminals and the office computers, the system has gained growing popularity for process attribute monitoring. The waste ratio for the whole division, calculated by the

Table 26 Summary of results for waste reduction and energy

Compare Type	Unit	2012 vs. 2014	2013 vs. 2014
		[Mean]	[Mean]
Waste Material	[%]	-	-0.54
Waste Material	[kg]	-	-54,400
Waste Rolls	[%]	-	-
Waste Rolls	[#]	-	-
Savings	[%]	-	-1.6
Savings	[EUR]	-	-400,000
Energy Savings	[EUR]	-5,600,000	-2,800,000

Controlling department as the main objective institution at the plant, decreased by 0.54% compared to 2013, Table 26. To make a final objective conclusion pertaining to all machines, data from about 1 year of implementation and continuous usage have to be analyzed.

In ongoing developments, eight further machines from the previous and the downstream sub-divisions, i.e. the compounders and the converter machines, respectively, will be integrated.

10 Summary and Conclusions

10.1 Summary

The ongoing industrial revolution due to Industrie 4.0 and Big Data incorporates the advance and massive employment of large scale data analysis by investigation of process data in manufacturing companies for process optimization, yield optimization, and sustainable use of resources [[Win13] [Bun15]].
Todays polymer film industry, serving as a research vehicle for the presented investigation, has a high time and cost spending effort in analyzing process problems, due to missing monitoring systems, unlinked data sources and a lack in explorative research. In particular, dynamic condition changes leading to material waste are not monitored from a holistic point of view, collectively regarding available distributed multi-sensory information at the same time.
Besides, the capabilities of *State-of-the-art* classification methods' functionalities are limited regarding typical one-class problems, caused by an overrepresentation of positive or fault-free examples whereas the number of definite negative or faulty examples is too low.
The explored analytical results are mostly presented at high level of research requirements and not understandable for the typical operator at the machine and related industrial requirements. Furthermore *State-of-the-art* standardized monitoring tools are not flexible enough for fast adaptation, whereas special additional extensions, as e.g., visualization, increase the development time and cost do not consider all needed functionalities or specifications.
The consensus between explorative research and industrial adaptation was focused on in this thesis. Therefore, adjustable methods for explorative investigation with integrative functionalities for an adaptive graphical user interface prototyping, applying the operators needs, and considering self-x properties are desired.
The thesis, based on the approach and goals outlined in Chapter 1, presents a design architecture for multi-sensory data analysis and on-line evaluation for advanced process monitoring with regard to industrial requirements, learning/ adapting from related fields, e.g. MEL.

The investigated topics are enumerated in the following:

- Knowledge-discovery-process for implementation into a dynamic open-loop integrated into a system architecture for projection on a research approach for basic conception of future self-x ideas.
- Optimizing the recognition behaviour of modified one-class classification methods on real datasets from the research vehicle.
- Visualization of high dimensional into lower dimensional space by multiple scaling approaches, as e.g., Sammon's Mapping.
- Investigation of Big Data methods for process trajectory behaviour in lower dimensional space for earlier condition classification improvement.
- Development of graphical human-machine-interface (HMI) for real-time monitoring & recommendation combining the research and industrial requirements.

The proposed system architecture for the real-time computer-aided manufacturing processing action & recommendation system has been partially implemented. The layer contains the acquisition and the monitoring & recommendation layer for information exchange between oracle database and machine PLC for analysis and further visualization. The adaptation layer consisting self-x properties for interaction of objects is still emerging and not yet implemented, due to readability of the PLC and restricted writeability.
An open-loop monitoring approach processes the acquired datasets with a feature reduction method to lower dimensions. The reduced datasets are then analyzed by an one-class classification approach which is adapted to training datasets. This is done by modification of the method settings. Each time a new dataset is acquired it is evaluated by the current one-class model to determine which class is best fitting, the class for normal or for abnormal conditions.
The visualization into lower dimensional space by comparing classical and non-classical methods achieved good results for both groups in 3- and 2-dimensional space, extendable to 1-dimensional with acceptable reconstruction error. The achieved mapped 2-dimensional view is separated into 3 zones by threshold limitation. The first zone for normal condition, the second boundary zone for conditions with trend to the third zone, and the third the abnormal condition zone.
The described view marks the basic for the following trajectory classification approach, that archieved by about 10 minutes faster results in advance compared to the OCC approach. By entering the boundary zone the following timestamps/minutes are observed. As far as a particular number of timestamps in sequence is achieved the trajectory is defined and evaluated by a classification approach. The previous OCC approach could be enhanced by the trajectory method. The advancement from the OCC to a trajectory approach supports carries the promise of faster recognition of abnormal conditions.
The development of domain-specific graphical user interface giving recommendations about occuring process condition deviations in real-time including interactive visualization techniques is essential for the acceptance of the approach. The open-loop behaviour is tracable and optimized for monitoring industrial requirements. Also, it is flexibly extendable. The returned classification results are visualizable and simplified for downstream analysis.

10.2 Novel Contributions

The, at least for polymer manufacturing industry, novel approach of multi-step process condition classification is combined with advanced visualization methods leading to earlier recognition of unknown states. The holistic design prepares the next stage for self-x optimization in future works.
This thesis focuses on the proposed system architecture depicted in Chapter 6, which is the basis for the compared approaches outlined in Chapter 7 and their implementation, as described in Chapters 8 and 9. The investigated novelties are explained following the sequence outlined below:

- Survey on industrial data processing requirements according to Industrie 4.0 and Big Data investigations
- Consideration of the industrial requirements to evolve an open-loop monitoring

approach with predictable behavior for specific critical problems

- First-time investigation of comprehensive manufacturing datasets in polymer film industry with modified *State-of-the-art* classification methods within an explorative knowledge discovery process
- Novel visualization into lower-dimensional space by multiple scaling approaches in 1-, 2-, or 3-dimensional view for laboratory and on-line applications
- Investigation of Big Data methods for the determination of trajectory behavior in lower-dimensional space for earlier condition classification improvement, with development of a novel trajectory extraction approach
- Development of a novel domain-specific graphical human-machine-interface (HMI) for real-time monitoring & recommendation combining the research and industrial requirements

The investigation of the manufacturing industry according to Industrie 4.0, and particularly the polymer film industry as a research vehicle, helped detect new room for improvement. Exploring Big Data in industrial plants with advanced methods enables faster reaction to occurring process deviations.
Typical process problems such as the "cut-off" problem were identified and aims were defined, e.g., lowering the material waste due to the specific issue by 5%. 21,900 datasets acquired and selected from 2 years of production were automatically labeled by a manufacturing execution system and then explored. The investigated classification methods, modified and assessed for one-class classification, achieved best results of 99% recognition rate due to their ability to handle classes with underrepresented samples, Table 14.
The result visualization of the one-class classification for downstream investigation was achieved with 2-dimensional scaling approaches. The classical and non-classical mapping approaches were assessed and compared. In this case, the Sammon's Mapping offers several advantages, such as the lowest reconstruction error in plotting new datasets into a predefined view, separated into different zones according to the problem condition or the normal production process.
Next, a – for polymer industry – novel approach to trajectory selection according to sliding windows and segmentation methods was developed on the above-mentioned scaled view. Through investigation of the zones and their subsequent new dataset positions within a prototype squared matrix, trajectories were extracted with downstream classification ratios of 99% from Table 22.
The described steps (one-class classification, Sammon's Mapping, and trajectory classification) were integrated into the specifically developed system architecture, applying data acquisition and monitoring & recommendation layers.
The complete setup was implemented in a first prototype approach as a domain-specific graphical user interface. It was tested and evaluated with real, exemplary manufacturing datasets to select the best-fitting method. Furthermore, the real-time test was carried out at a laboratory prototype machine and, after successful investigation, was extended to a group of on-line production machines. The interface was appropriately upgraded and adapted to the specific environment.
So far, since its installation and launch, the implemented system has processed more than 200 million datasets of the laboratory machine, with a system precision of 90% of

correctly recognized "cut-off" problems. The product quality has improved, expressed by at least 2% decrease of the critical material waste problem related to 50,000 EUR savings per month from 2013 to 2014 at the machine M150. The complete material waste of the machine decreased from 20,96% auf 12,44% within the last two years.
The complete approach cycle focusing on the settings for the knowledge discovery processes is presented in Fig. 99.

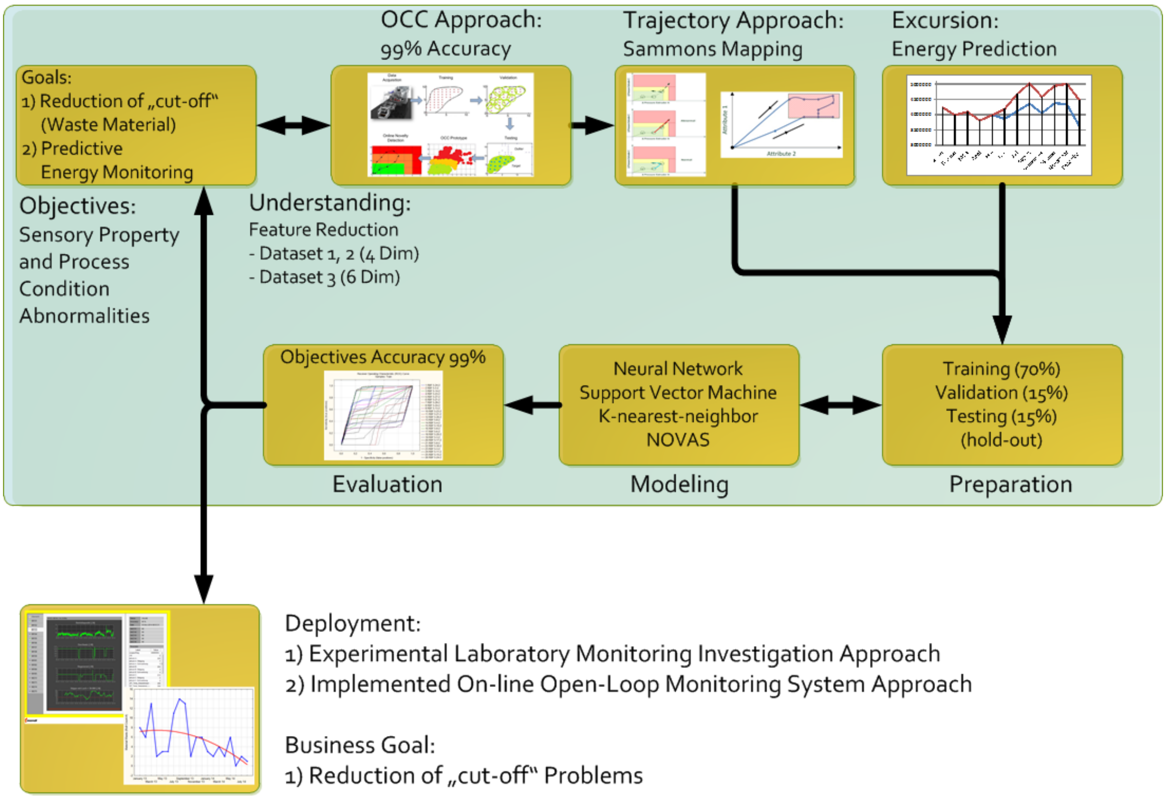


Figure 99 Online, and laboratory open-loop monitoring system research approach according to the knowledge discovery cycle (CRISP)

The following discussion reflects the advantages and disadvantages of the current approach.
The material waste reduction as the main target was achieved, although only one specific problem was regarded and solved. The investigation should be extended to more particular critical cases from the polymer production process, to demonstrate the general validity of the system in predicting the occurrence of unknown conditions.
Regarding the sensor data points, the acquired information was not tested with regard to sensor drift effects and quality degradation. The diagnostic procedures were not investigated in this work; such an investigation could prevent system failures due to low data quality.
The restriction to 6 dimensions out of 160 at an early stage of the investigation is questionable due to the limited time range of the acquired datasets for feature reduction.
The recognition rate depends on the datasets used for training; therefore, a cyclical training loop verification should perhaps be additionally integrated in order to recalibrate and dynamically adapt the complete system architecture by extension of the training cases.
Further on, the threshold setting for the boundary zone does not exactly display the

class separation in the regarded dimension. Therefore, the trajectory acquisition by equal squared segmentation without investigation of other possible approaches could weaken the result stability.
The acquisition rate in 1-min intervals was not investigated due to changes in recognition behavior depending on the varying refreshing rates. Lowering the rate from, e.g., 1- to 10-min intervals could decrease the storage amount, thus relieving the database combined with cost reductions.
The human-machine interface was individually adapted to the particular manufacturing process without benchmarking other similar industry types in monitoring approaches, as e.g., the paper mills or nonwoven production plants.

Regarding the research approaches in the thesis, parts of the results are published in the following publications [[KK12], [KK15]], and presentation on the 6th OCS Global Gel Meeting [Koh13] novel contributions were achieved in parts of OCC and trajectory investigation.

10.3 Future Work

In the following years, the usage in industrial plants of advanced sensor-based electronical systems for, e.g., process recognition and self-adjustment will move into the focus in connection with Industrie 4.0 [BMB14].
Further research on mapping will be performed and novel classification-based selection techniques will be investigated. The recognition ratio from the current classification approaches strongly depends on the given datasets. Therefore, more datasets from a wider time range and from different objects have to be analyzed and compared in further investigations.
In further developments, the spectrum of methods will be extended and sensor failures will be analyzed in more detail with, for instance, diagnostic approaches. These diagnostic approaches offer essential improvements for the recognition of sensor failures.
The researched methods will be extended to other machines as, e.g., the blown film machines, and the examined results will be improved for implementation in so-called off-line and on-line recognition systems for polymer process monitoring, as presented in Chapters 8 and 9. Four converter and four compounder machines shall be integrated into the online open-loop control for condition monitoring of the whole three-stage elastomer process of the current research vehicle, as described in Figure 11 (see Chapter 3).
Process behavior simulations will be introduced to develop suggestions within virtual environments[37].
The proposed system architecture part for self-x properties will be activated for the self-adjustment of manufacturing machines. Machine settings will be programmed and adjusted automatically by the monitoring program for earlier reaction. The developments made for the prediction of energy consumption will be further investigated to achieve higher precision, and a plant forecast will be examined, combining the customer, manufacturer and delivery stages, as displayed in Fig. 100.

[37] JMP, part of SAS for interactive data visualization and analysis

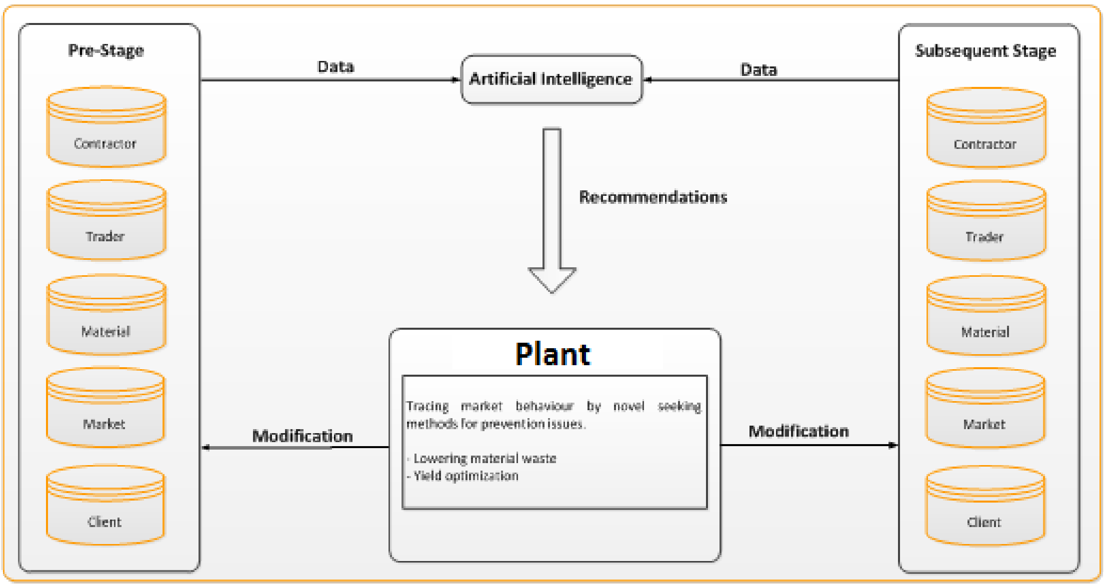


Figure 100 Outlook on new predictive project for the plant forecast

In the future, companies will continuously analyze the acquired and merged *Big Data* types from the production process, the division, and the complete plant, giving recommendations and acting in a self-adjusting way. Therefore, data scientists and server admins will become more important for such companies, to build up explorative knowledge discovery approaches in huge databases [$S^{+}12$].

Index of Abbreviations

ABS .... Acrylonitrile butadiene styrene

ASA .... Acrylonitrile styrene acrylate

CAQ .... Computer Assisted Quality Control

CNC .... Computerized Numerical Control

ConMon ..... Condition Monitoring

CPPS .... Cyber-Physical Production System

CRISP .... Cross-Industry Standard Process

EEG .... Renewable Energy Law

EKZ .... Energy Ratio

GUI .... Graphical User Interface

HMI .... Human-Machine-Interface

IoT .... Internet of Things

KDP .... Knowledge-Discovery-Process

KNN .... Nearest-Neighbor Classifier

kW .... Kilowatt

kWh .... Kilowatt Hour

MDE .... Machine-Data-Acquisition

MDS .... Multidimensional Scaling

MEL .... Microelectronics

MES .... Manufacturing Execution System

MLP .... Multilayer Perceptron

NLMR .... Non-Linear Mapping Recall

NOVAS ...... Novelty Associative Memory

NOCLASS ... Novelty Classification

OCC .... One-Class Classification

OCS .... Optical Control System

OECD .... Organization for Economic Co-operation and Development

PA .... Polyamide

PC .... Polycarbonate

PE .... Polyethylene

PE-HD .... Polyethylene, high density

PE-LD .... Polyethylene, low density

PE-LLD ...... Polyethylene, linear low density

PE-MD ...... Polyethylene, medium density

PEMRG ..... PlasticsEurope Market Research Group

PET .... Polyethylene terephthalate

PKZ .... Process Ratio

PLC .... Programmable Logic Controler

PMMA .... Polymethyl methacrylate

PP .... Polypropylene

PRF .... Plastics Recovery Facilities

PS .... Polystyrene

PS-E .... Polystyrene, expandable

PUR .... Polyurethane

PVC .... Polyvinyl chloride

QKZ .... Quality Ratio

RBF .... Radial Basis Function

R&D .... Research & Development

RNN .... Reducing-Nearest-Neighbor Classifier

ROC .... Receiver Operating Characteristic

SAN .... Styrene-acrylonitrile

SBS .... Styrol-blockcopolymere

sbs .... Sequential Backward Selection

Self-X .... Self-* Properties

sfs .... Sequential Forward Selection

SOM .... Self Organizing Maps

SPC .... Statistical Process Control

SQL .... Structured Query Language

SVM .... Support Vector Machine

VLAN .... Internal Machine Network

XML .... Extensible Markup Language

List of Symbols

C .... Covariance Matrix

d .... Distance

D .... Distance Matrix

D_{nov}^{ij} .... Novelty Distance

err_{ij} .... Error Function

E .... Sammon's Error

G_{nov}^{ij} .... Grey-Level Value

IG .... Information Gain Ratio

$\vec{p}_{ij}$ .... Pixel Vector

Rej .... Rejection Ratio

S .... Training Sample

V .... n-by-n Diagonal Matrix

w_{ij}^{NN} .... Neighbor Neuron

$\bar{x}$ .... Mean Value

X_{ij} .... Data Sample

x_i, c_i .... Instance

x_{ki} .... Attribute

y .... Observation

Y_{ij} .... Pivot Point

Θ_{ij} .... Threshold

References

[Abe12] Aberdeen Group. Big Data for Small Budgets. Technical Report, Aberdeen Group, Boston, 2012.

[Alt92] N. S. Altmann. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *The American Statistician*, Vol.46(3):175–185, 1992.

[B+11] O. Boehm et al. Classifying Cognitive States of Brain Activity via One-Class Neural Networks with Feature Selection by Genetic Algorithms. *International Journal of Machine Learning and Cybernetics*, Vol.2(3):125–134, 2011.

[Bai94] K. D. Bailey. *Numerical Taxonomy and Cluster Analysis.* SAGE Publications Inc., Thousand Oaks, 1994.

[Ban13] A. Banerjee. Big Data & Advanced Analytics in Telecom: A Multi-Billion-Dollar Revenue Opportunity. Technical Report, Heavy Reading, New York, 2013.

[Bar10] A. Bartkowiak. Anomaly, Novelty, One-Class Classification: A Short Introduction. *Computer Information Systems and Industrial Management Applications (CISIM)*, pp.:1–6, 2010.

[BG05] I. Borg and P. J. F. Groenen. *Modern Multidimensional Scaling: Theory and Applications*. Springer, New York, 2005.

[Bha05] P. Bhagat. *Pattern Recognition in Industry.* Elsevier Science, London, 2005.

[Bis95] C. M. Bishop. *Neural Networks for Pattern Recognition.* Oxford University Press Inc., New York, 1995.

[BMB14] BMBF. Sensorbasierte Elektroniksysteme für Anwendungen für Industrie 4.0 (SElekt I4.0). http://www.bmbf.de/foerderungen/25379.php, 2014. Visited on 01/07/15.

[Bor09] K. Borne. Scientific Data Mining in Astronomy. *Next Generation of Data Mining*, pages 91–114, 2009.

[Bra02] D. Braha. *Data Mining for Design and Manufacturing: Methods and Applications (Massive Computing).* Springer, London, 2002.

[BtH07] H.-J. Bullinger and M. ten Hompel. *Internet der Dinge.* Springer, Berlin, 2007.

[Bun13] Bundesnetzagentur. Netzentgelte. http://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/Netzentgelte/netzentgelte-node.html, 2013. Visited on 01/07/15.

[Bun14] Bundesamt für Wirtschaft und Ausfuhrkontrolle. Besondere Ausgleichsregelung §§ 63 ff. EEG. http://www.bafa.de/bafa/de/, 2014. Visited on 01/07/15.

[Bun15] Bundesministerium für Wirtschaft und Energie. Autonomik für Industrie 4.0. http://www.autonomik.de/de/1003.php, 2015. Visited on 01/07/15.

[Bur98] C. J. C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, Vol.2:121–167, 1998.

[Bus15] Busitech. Firm Description. http://www.busitech.com/, 2015. Visited on 01/07/15.

[C+04] C. K. Cheng et al. Enhancing Neural Confidence-Based Segmentation for Cursive Handwriting Recognition. *Proceeding of 5th International Conference on Simulated Evolution and Learning (SEAL'04)*, pages SWA–8, 2004.

[C+07] K. J. Cios et al. *Data Mining: A Knowledge Discovery Approach*. Springer, London, 2007.

[CM95] M. Costa and L. Moura. Automatic Assessment of Scintmammographic Images using a Novelty Filter. *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, Vol.1:537–541, 1995.

[CST99] N. Cristianini and J. Shawe-Taylor. *An Introduction to Support Vector Machines: and other Kernel-Based Learning Methods*. Cambridge University Press, New York, 1999.

[CST00] N. Cristianini and J. Shawe-Taylor. *Support Vector Machines*. Cambridge University Press, New York, 2000.

[CU93] A. Cichoski and R. Unbehauen. *Neural Networks for Optimization and Signal Processing*. John Wiley & Sons, Hoboken, 1993.

[CV95] C. Cortes and V. Vapnik. Support-Vector Networks. *Machine Learning*, Vol.20(3):273–297, 1995.

[D+04] K.-H. Diener et al. Modellierung und Simulation zur Entwurfsunterstützung integrierter piezoresistiver Druckmesssysteme. Technical Report, Fraunhofer-Institut für Integrierte Schaltungen, Dresden, 2004.

[D+07] R. P. W. Duin et al. A Matlab Toolbox for Pattern Recognition. PRTools. http://prtools.org/, 2007. Visited on 01/07/15.

[D+10] A. Depeursinge et al. Comparative Performance Analysis of State-Of-The-Art Classification Algorithms Applied to Lung Tissue Categorization. *Journal of Digital Imaging*, Vol.23(1):18–30, 2010.

[D+12] H. Domininghaus et al. *Kunststoffe*. Springer, Heidelberg, 2012.

[Dan05] T. L. Daniel. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley & Sons, Hoboken, 2005.

[Dav14] T. Davenport. A Predictive Analytics Primer. http://blogs.hbr.org/2014/09/a-predictive-analytics-primer/, 2014. Visited on 01/07/15.

[Del34] B. N. Delaunay. Sur la sphere vide. *Bulletin of Academy of Sciences of the USSR*, Vol.6:793–800, 1934.

[Dex06] L.A. Dexter. *Elite and Specialized Interviewing*. E C Prod Inc., Houston, 2006.

[DW92] C. E. Dunn and P. S. P. Wang. Character Segmentation Techniques for Handwritten Text-A Survey. *Proceeding of International Conference on Pattern Recognition*, pages 577–580, 1992.

[F+96] L. Fahrmeir et al. *Multivariate statistische Verfahren*. de Gruyter, Berlin, 1996.

[F+13] J. Fähndrich et al. Towards Self-Explaining Agents. *Trends in Practical Applications of Agents and Multiagent Systems*, Vol.221:147–154, 2013.

[Fay96] U. M. Fayyad. Data Mining and Knowledge Discovery in Databases: Applications in Astronomy and Planetary Science. *Proceedings of the Thirteenth National Conference on Artificial Intelligence and Eighth Innovative Applications of Artificial Intelligence Conference, AAAI-96 Proceedings*, Vol.2:673–684, 1996.

[G+13] O. Ganschar et al. Produktionsarbeit der Zukunft – Industrie 4.0. Technical Report, Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Stuttgart, 2013.

[Gla14] A. Glaser. What Exactly is the Internet of Things? Technical Report, Harbor Research, Boulder, 2014.

[Gna97] R. Gnanadesikan. *Methods for Statistical Data Analysis of Multivariate Observations*. John Wiley & Sons, Hoboken, 1997.

[Gne14] Gneuss Kunststofftechnik GmbH. Firm Description. http://www.gneuss.de/, 2014. Visited on 01/07/15.

[H+01] D. Hand et al. *Principles of Data Mining*. MIT Press, Cambridge, 2001.

[H+05] J. A. Harding et al. Data Mining in Manufacturing: A Review. *Journal of Manufacturing Science and Engineering*, Vol.128(4):969–976, 2005.

[H+09] T. Hastie et al. *The Elements of Statistical Learning*. Springer, London, 2009.

[H+14] J. Holler et al. *From Machine-to-Machine to the Internet of Things: Introduction to a New Age of Intelligence*. Academic Press, London, 2014.

[HA04] V. Hodge and J. Austin. A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review*, 22(2):85–126, 2004.

[Hay98] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall PTR, Upper Saddle River, 1998.

[HP13] M. Hadhri and J. Pieper. Cefic – The European Chemicals Industry. Technical Report, Cefic Chemdata International, Brussels, 2013.

[HvS00] A. Heim and O. von Stryk. Trajektory Optimization of Industrial Robots with Application to Computer-Aided Robotics and Robot Controllers. *Journal of Digital Imaging*, Vol.47:407–420, 2000.

[JCG08] JCGM. Evaluation of measurement data - Guide to the expression of uncertainty in measurement. Technical Report, JCGM member organizations, 2008.

[JN02] L. B. Jack and A. K. Nandi. Fault Detection using Support Vector Machines and Artificial Neural Networks, Augmented by Genetic Algorithms. *Mechanical Systems and Signal Processing*, Vol.16(2–3):373–390, 2002.

[K+92] C. Kohlert et al. *Kalandrieren von Polymeren.* Deutscher Verlag für Grundstoffindustrie, Leipzig, 1992.

[K+93] A. Kaltenmeier et al. *Hidden Markov Models – A Unified Approach to Recognition of Spoken and Written Language*, chapter Mustererkennung, 15. DAGM Symposium, pages 191–198. Springer, Berlin, 1993.

[K+94a] A. König et al. A Novel Method for the Design of Radial-Basis-Function-Networks and its Implication for Knowledge Extraction. *In Proceedings of the International Conference on Neural Networks ICNN'94*, Vol.3:1804–1809, Orlando, Florida, 1994.

[K+94b] A. König et al. Massively Parallel VLSI-Implementation of a Dedicated Neural Network for Anomaly Detection in Automated Visual Quality Control. *In Proceedings of the 4th Intern. Conf. on Microelectronics for Neural Networks and Fuzzy Systems*, pp.:354–363, Torino, Italy, 1994.

[K+95] A. König et al. Visual Inspection in Industrial Manufacturing. *IEEE Microelectronics for Neural Networks and Fuzzy Systems*, Vol.15:26–31, 1995.

[K+98] A. König et al. A Transparent and Flexible Development Environment for Rapid Design of Cognitive Systems. *Euromicro Conference*, pp.:655–662, 1998.

[Kö00] A. König. Interactive Visualisation and Analysis of Hierarchical Neural Projections for Data Mining. *IEEE Transactions on Neural Networks TNN*, Vol.11(3):615–624, 2000.

[Kö01] A. König. *Dimensionality Reduction Techniques for Interactive Visualisation, Exploratory Data Analysis and Classification*, chapter Pattern Recognition in Soft Computing Paradigm, FLSI Soft Computing Series, pages 1–37. World Scientific Publishing Co. Inc., River Edge, 2001.

[K+10] M. Kohlert et al. *Automated Anticounterfeiting Inspection Methods for Rigid Films Based on Infrared and Ultraviolet Pigments and Supervised Image Segmentation and Classification*, chapter Knowledge-Based and Intelligent Information and Engineering Systems, pages 321–330. Springer, Heidelberg, 2010.

[K+13] H. Kagermann et al. Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Technical Report, Arbeitskreis Industrie 4.0, Frankfurt, 2013.

[KC03] J. O. Kephart and D. M. Chess. The Vision of Autonomic Computing. *Computer*, Vol.36:41–50, 2003.

[KDN12] KDNuggets. Analytics, Data Mining, Big Data Software Used. http://www.kdnuggets.com/polls/2012/analytics-data-mining-big-data-software.html, 2012. Visited on 01/07/15.

[KG05] A. König and A. Gratz. *Advanced Methods for the Analysis of Semiconductor Manufacturing Process Data*, chapter Advanced Techniques in Knowledge Discovery and Data Mining, pages 27–74. Springer, Heidelberg, 2005.

[KK12] M. Kohlert and A. König. Advanced Polymeric Film Production Data Analysis and Process Optimization by Clustering and Classification Methods. *Frontiers in Artificial Intelligence and Applications*, Vol.243:1953–1961, 2012.

[KK15] M. Kohlert and A. König. Large, High Dimensional, Heterogeneous Multi-Sensor Data Analysis Approach for Process Yield Optimization in Polymer Film Industry. *Neural Computing and Applications*, Vol.26(3):581–588, 2015.

[KM10] S. Khan and M. G. Madden. A Survey of Recent Trends in One Class Classification. *AICS'09 Proceedings of the 20th Irish Conference on Artificial intelligence and Cognitive Science*, Vol.1:188–197, 2010.

[Koh13] M. Kohlert. Web Inspection Systems for Advanced Film Production: 6th OCS' Global Gel Meeting. Conference Presentation, 2013.

[Kru64a] J. B. Kruskal. Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis. *Psychometrika*, Vol.29(1):1–27, 1964.

[Kru64b] J. B. Kruskal. Nonmetric Multidimensional Scaling: A Numerical Method. *Psychometrika*, Vol.29(2):115–129, 1964.

[Kue99] B. Kueppers. *Data Mining in der Praxis: Ein Ansatz zur Nutzung der Potentiale von Data Mining im betrieblichen Umfeld*. Lang, Frankfurt, 1999.

[Kus06] A. Kusiak. Data Mining: Manufacturing and Service Applications. *International Journal of Production Research*, Vol.44(18):4175–4191, 2006.

[KW78] J. B. Kruskal and M. Wish. *Multidimensional Scaling*. SAGE, London, 1978.

[L+08] G. Lee et al. Trajectory Outlier Detection: A Partition-and-Detect Framework. *Proceedings of the 24th International Conference on Data Engineering*, Vol.24:140–149, 2008.

[M. 02] M. Loughran. IBM Unveils New Autonomic Computing Deployment Model. http://www-03.ibm.com/press/us/en/pressrelease/464.wss, 2002. Visited on 01/07/15.

[M+01] J. Moyne et al. *Run-to-Run Control in Semiconductor Manufacturing*. CRC Press LLC, Boca Raton, 2001.

[M+14] T. McGuire et al. Why Big Data is the New Competitive Advantage. http://iveybusinessjournal.com/publication/why-big-data- is-the-new-competitive-advantage/, 2014. Visited on 01/07/15.

[Mar08] W. Martin. Trends in Data Mining. http://www.beyenetwork.de/view/8192/, 2008. Visited on 01/07/15.

[Mar11] W. Martin. Trends in Data Mining. http://www.beyenetwork.de/view/14993/, 2011. Visited on 01/07/15.

[Mar12] W. Martin. Trends in Data Mining. http://www.beyenetwork.de/view/15931/, 2012. Visited on 01/07/15.

[MH96] M. M. Moya and D. R. Hush. Network Constraints and Multi-Objective Optimization for One-Class Classification. *Neural Networks*, Vol.9(3):463–474, 1996.

[Mit97] T. Mitchel. *Machine Learning*. McGraw-Hill, New York, 1997.

[Mon15] Mondi Gronau GmbH. Firm Description. http://www.mondigroup.com/, 2015. Visited on 01/07/15.

[MY01] L. M. Manevitz and M. Yousef. One-Class SVMs for Document Classification. *Journal of Machine Learning*, Vol.2:139–154, 2001.

[Nec10] P. Neckel. mayato Data Mining Studie 2010. Technical Report, mayato, Berlin, 2010.

[OCS15] OCS Optical Control Systems GmbH. Firm Description. http://www.ocsgmbh.com/, 2015. Visited on 01/07/15.

[OEC14] OECD. Perspective on Global Developments 2014. Technical Report, Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Stuttgart, 2014.

[Ora14a] Oracle Corporation. Oracle Database. http://www.oracle.com/de/database/overview/index.html, 2014. Visited on 01/07/15.

[Ora14b] Oracle Corporation. Oracle Database Concepts. http://docs.oracle.com/cd/E11882_01/server.112/e25494.pdf, 2014. Visited on 01/07/15.

[P+03] A. Padovitz et al. Awareness and Agility for Autonomic Distributed Systems: Platform-Independent Publish-Subscribe Event-Based Communication for Mobile Agents. *Database and Expert Systems Applications*, pp.:669–673, 2003.

[Per07] M. Perzyk. Statistical and Visualization Data Mining Tools for Foundry Production. *Archives of Foundry Engineering*, Vol.7(3):111–116, 2007.

[Pet12] C. Pettey. Gartner Says Big Data Creates Big Jobs: 4.4 Million IT Jobs Globally to Support Big Data by 2015. http://www.gartner.com/newsroom/id/2207915, 2012. Visited on 01/07/15.

[PJ05] N. R. Pal and L. C. Jain. *Advanced Techniques in Knowledge Discovery and Data Mining*. Springer, London, 2005.

[Pla13] Plastics Europe. Plastics – the Facts 2013. Technical Report, Plastics Europe, Frankfurt, 2013.

[Qui86] J. R. Quinlan. Introduction of Decision Trees. *Machine Learning*, Vol.1(1):81–106, 1986.

[R+98] D. Ridder et al. An Experimental Comparison of One-Class Classification Methods. Conference Paper. *ASCI*, pp.:213–218, 1998.

[RD97] D. Ridder and R.P.W. Duin. Sammon's mapping using neural networks: A comparison. *Pattern Recognition Letters*, Vol.18:1307–1316, 1997.

[Rex13] Rexer Analytics. Data-Mining-Anwenderstudie 2013 von Rexer Analytics. Technical Report, Rexer Analytics, Winchester, 2013.

[RG07] A. Reyes and D. Gilbert. Prediction of Protein-Protein Interactions using One-Class Classification Methods and Integrating Diverse Biological Data. *Journal of Integrative Bioinformatics*, Vol.4(3):77, 2007.

[S+99] B. Schölkopf et al. Support Vector Method for Novelty Detection. *NIPS*, Vol.12:582–588, 1999.

[S+05a] H. J. Shin et al. One-Class Support Vector Machines: An Application in Machine Fault Detection and Classification. *Computers & Industrial Engineering*, Vol.48(2):395–408, 2005.

[S+05b] I. Steinwart et al. A Classification Framework for Anomaly Detection. *The Journal of Machine Learning Research*, Vol.6:211–232, 2005.

[S+05c] D. Sun et al. A Novel Method for Chinese Spam Detection Based on One-Class Support Vector Machine. *Information and Computational Science*, Vol.2(1):109–114, 2005.

[S+12] M. Sinn et al. Data Science – Die Karriere der Zukunft. Technical Report, EUROFORUM, Düsseldorf, 2012.

[S+14] G. Spur et al. *Handbuch Urformen.* Carl Hanser, München, 2014.

[Sag12] E. Sagel. Polyethylene Global Overview. Technical Report, Pemex and IHS, Mexico City, 2012.

[Sam69] J. W. Sammon. A Nonlinear Mapping for Data Structure Analysis. *IEEE Transactions on Computers*, Vol.18(5):401–409, 1969.

[SAS14] SAS Institute Inc. Mehr Erfolg in der Kundenakquise, besserer Service: Das Karriereportal Monster Canada setzt auf SAS. http://www.sas.com/de_de/customers/monster-canada.html, 2014. Visited on 01/07/15.

[Sta14a] Statistisches Bundesamt and VCI. Spartenumsätze in der chemisch-pharmazeutischen Industrie in Deutschland. http://de.statista.com/statistik-/daten/studie/206253/umfrage/spartenumsaetze-in-der-chemisch-pharmazeutischen-industrie-in-deutschland/, 2014. Visited on 01/07/15.

[Sta14b] Statistisches Bundesamt and VCI. Umsätze der wichtigsten Industriebranchen in Deutschland. http://de.statista.com/statistik/daten/studie/241480/umfrage/umsaetze-der-wichtigsten-industriebranchen-in-deutschland/, 2014. Visited on 01/07/15.

[Sta14c] StatSoft GmbH. Access SQL: Grundkonzepte, Sprache und Syntax. http://office.microsoft.com/de-de/access-help/access-sql-grundkonzepte-sprache-und-syntax-HA010256402.aspx, 2014. Visited on 01/07/15.

[Ste09] J. Stevens. *Applied Multivariate Statistics for the Social Science*. L. Erlbaum Associates Inc., Hillsdale, 2009.

[Stu11] E. Stubbs. The Value of Business Analytics. http://www.analytics-magazine.org/special-articles/461-the-value-of-business-analytics, 2011. Visited on 01/07/15.

[SV11] K. Smets and J. Vreeken. The Odd One Out: Identifying and Characterising Anomalies. *SDM'11*, pp.:804–815, 2011.

[T+95] L. Tarassenko et al. Novelty Detection for the Identification of Masses in Mammograms. *Proceedings of the 4th International Conference on Artificial Neural Networks*, Vol.4:442–447, 1995.

[Tax01] D. M. J. Tax. *One-Class Classification: Concept-Learning in the Absence of Counter-Examples*. PhD thesis, University of Delft, 2001.

[Tax13] D. M. J. Tax. Data Description Toolbox. Pattern Recognition Laboratory. http://prlab.tudelft.nl/david-tax/dd-tools.html, 2013. Visited on 01/07/15.

[TD04] D. M. J. Tax and R. P. W. Duin. Support Vector Data Description. *Machine Learning*, Vol.54(1):45–66, 2004.

[The15] The MathWorks Inc. Firm Description. http://de.mathworks.com/, 2015. Visited on 01/07/15.

[Tor52] W. S. Torgerson. Multidimensional Scaling: I. Theory and Method. *Psychometrika*, Vol.17(4):401–419, 1952.

[Tor58] W. S. Torgerson. *Theory and Methods of Scaling*. John Wiley & Sons, New York, 1958.

[Twe14] D. Tweed. Science, Models and Machine Learning. http://johncarlosbaez.wordpress.com/2014/09/03/science-models-and-machine-learning/, 2014. Visited on 01/07/15.

[V+02] G. Vlachos et al. Discovering Similar Multidimensional Trajectories. *Proceedings of the International Conference on Data Engineering*, pp.:673–684, 2002.

[V+09] VDI et al. Prozess-Sensoren 2015+. Technical Report, VDI, VDE, NAMUR, Leverkusen, 2009.

[VS14] S. Vongsingthong and S. Smanchat. Internet of Things: A Review of Applications & Technologies. *Suranaree Journal of Science and Technology*, pp.:1–24, 2014.

[W+94] H. Wapenhans et al. Optimal Trajectory Planning with Application to Industrial Robots. *The International Journal of Advanced Manufacturing Technology*, Vol.9(1):49–55, 1994.

[W+08] R.P. Würtz et al. *Organic Computing*. Springer, Heidelberg, 2008.

[Wan07] K. Wang. Applying Data Mining to Manufacturing: The Nature and Implications. *Journal of Intelligent Manufacturing*, Vol.18(4):487–495, 2007.

[Win13] J. Winterhagen. Revolution in den Werkhallen. http://www.faz.net/aktuell/, April 2013. Visited on 01/07/15.

[Zac12] M. Zacher. Big Data Analytics in Deutschland. Technical Report, IDC Manufacturing Insights, Framingham, 2012.

[Zil14] A. Zilch. Industrie 4.0 in Deutschland. Technical Report, Experton Group, Ismaning, 2014.

[ZW14] A. Zilch and M. Weiss. Große Anwenderstudie zu Industrie 4.0 in Deutschland. Technical Report, Experton Group, Ismaning, 2014.

Appendix A

Neural Networks Scaling Approach

Therefor a more representative iteration sequence was generated as a parallel approach with a more suitable stepwise overview for a better illustration of the iteration steps to the reader. According to the code by Gianniotis (2013) [The15], an ensemble of neural networks was implemented, which used the Sammon's Mapping metric for the iteration from high-dimensional into the lower wanted 2-dimensional space, Fig. 101. The ensemble parameters were defined by 50 neural networks, each containing 7 hidden neurons, a maximum of 200 iterations, and the early stopping depending on the error, to reduce the dimensionality to 2 dimensions.

In Figure 101 the multi- dimensional unscaled dataset projections formed to the 2-dimensional scaled view for interface plotting for the staff for easier understanding with Sammon's metric by neural networks [RD97] [Kö00] is presented. The scaling procedure was executed with an ensemble of 50 neural networks, each with 7 hidden neurons. The iterations count was limited to 200 for reducing the dimensionality of 21,900 training datasets. The iteration steps, visualized in Matlab, show the development of the dataset constellation from the origin 6-dimensional view, first plot, to the exemplary chosen 5th iteration, followed by the final 200th iteration, separated into zones by delaunay/ convex hull and threshold setting afterwards, according to the Fig. 56.

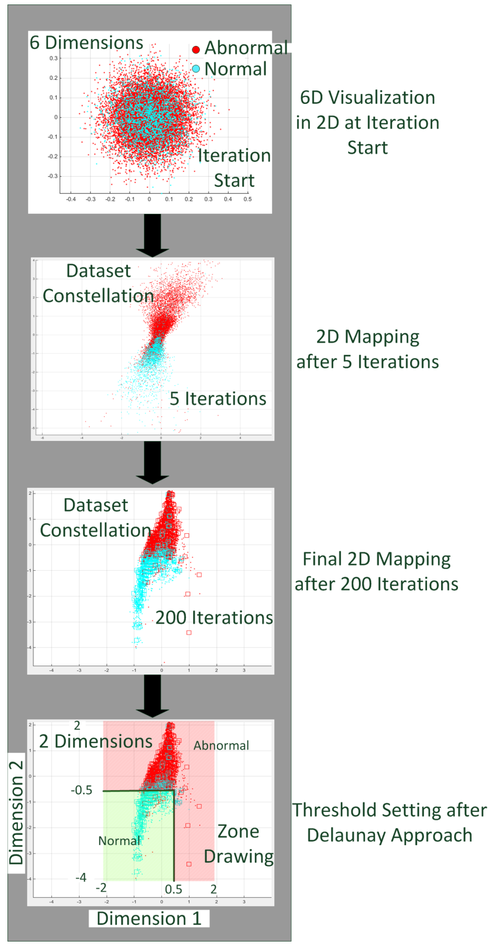


Figure 101 Multi-dimensional Scaling Procedure for Dimensionality Reduction from 6D to 2D with the Parallel Approach, the Ensemble of 50 Neural Networks for Projection

CURRICULUM VITAE

MICHAEL KOHLERT

EDUCATION

Technical University of Kaiserslautern *March 2010*
Diplom-Wirtschaftsingenieur Elektrotechnik

Thesis: Automated Anticounterfeiting Inspection Methods for Rigid Films based on Infrared and Ultraviolet Pigments and Supervised Image Segmentation and Classification

EXPERIENCE

Technical University of Kaiserslautern October 2012 - Present
External PhD *Kaiserlautern, RP*

- Integrated Sensor Systems, Department of Electrical Engineering and Computer Engineering
- Doctoral Thesis: Multi-Sensory Data Analysis and On-Line Evaluation for Advanced Process Control and Yield Optimization in Polymer Film Industry

Mondi Gronau GmbH March 2015 - Present
Head of Information Management & Process Automation *Gronau, NRW*

- IT Department

Mondi Gronau GmbH November 2014 - February 2015
Head of Process Automation *Gronau, NRW*

- Automation Department

Mondi Gronau GmbH April 2011 - October 2014
Assistant to Head of Production *Gronau, NRW*

- Production Department

Klöckner Pentaplast GmbH & Co. KG June 2010 - March 2011
Trainee *Montabaur, RP*

- Research & Development