DOKTORI (PhD) ÉRTEKEZÉS

DOBOS LÁSZLÓ

Pannon Egyetem

2015.

Kísérlettervezési technikák a technológiák elemzésére és optimálására

Értekezés doktori (PhD) fokozat elnyerése érdekében a
Pannon Egyetem Vegyészmérnöki Tudományok és Anyagtudományok
Doktori Iskolájához tartozóan

Írta:
Dobos László

Konzulens: Dr. Abonyi János
Elfogadásra javaslom (igen/nem) .....
(aláírás)

A jelölt a doktori szigorlaton ..... %-ot ért el.

Az értekezés bírálóként elfogadásra javaslom:
Bíráló neve: igen/nem .....
(aláírás)
Bíráló neve: igen/nem .....
(aláírás)

A jelölt az értekezés nyilvános vitáján .....%-ot ért el.
Veszprém, ...
a Bíráló Bizottság elnöke

A doktori (PhD) oklevél minősítése .....

.....
az EDT elnöke

Pannon Egyetem
Vegyészmérnöki és Folyamatmérnöki Intézet
Folyamatmérnöki Intézeti Tanszék

Kísérlettervezési technikák a technológiák elemzésére és optimálására

DOI: 10.18136/PE.2016.617

DOKTORI (PhD) ÉRTEKEZÉS

Dobos László

Konzulens
Dr. habil. Abonyi János, egyetemi tanár

Vegyészmérnöki Tudományok és Anyagtudományok Doktori Iskola
2015.

University of Pannonia
Faculty of Chemical and Process Engineering
Department of Process Engineering

Experimental Design Techniques for Analyzing and Optimization of Technologies

PhD Thesis

László Dobos

Supervisor
János Abonyi, PhD, regius professor

Doctoral School in Chemical Engineering and Material Sciences
2015.

Köszönetnyilvánítás

A dolgozat elkészültével az életem egy fontos szakasza is befejeződik. Ezen időszak alatt nem csupán szakmailag fejlődtem, hanem bővült a látó- és érdeklődési köröm egyaránt. Megtanultam, hogy a különböző elméletek leegyszerűsítése az egyik legnehezebb feladatok egyike, de feltétele annak a széleskörben alkalmazható tudásnak, ami nem kizárólag a mérnöki, hanem az élet egyéb területein is sikerrel alkalmazható. Ebben nagy szerepe volt témavezetőmnek, Dr. Abonyi Jánosnak, aki fáradhatatlanul egyensúlyozott a mérnöki és gazdasági területek határmezsgyéjén, bemutatva, hogy a különböző tudományterületek közti kapcsolatot az alapelvek egyezősége és kompatibilitása nyújtja, ahol az innováció az alapok újszerű értelmezésében rejlik. Ennek eredményeként nem a klasszikus vegyipari folyamatmérnöki problémákra fokuszáltunk a közös munka során, hanem olyan módszerek fejlesztésére, amelyek multidiszciplináris módon értelmezhetők. Köszönetet szeretnék mondani a tanszéki munkaközösségnek, akik mindeme munka közben gyakorlatias szemmel és konstruktívan egyengették az utamat. Köszönetet szeretnék mondani a családomnak és mindazoknak, akik biztos hátteret adtak akkor is, amikor a bizonyos hullámvölgyek alján voltam és akkor is, amikor együtt tudtunk örülni egy-egy sikernek, cikknek, sikeres közös munkának.

Kivonat

Kísérlettervezési technikák technológiák elemzésére és optimálására

Napjainkban a vegyipari technológiák üzemeltetése során a korszerű folyamatirányító rendszerek egyik feladata a folyamatadatok naplózása, aminek során elképesztő mennyiségű információ bújik meg a tárolt adatokban, ún. idősorokban. A folyamatadatok felhasználásában rejlő lehetőségeket az ipari gyakorlat csupán az elmúlt években kezdte el felhasználni az egyes üzemek, üzemrészek fejlesztése, során. Ezek célja a gazdasági haszon növelése, miközben egyre közelebb kerül az adott technológia a fizikai, kémiai törvényszerűségek szabta határaihoz. A legkorszerűbb hibadetektálási, folyamatszabályozási- és optimalizási megoldások matematikai modelleket használnak fel előrejelzésre, ezek további fejlesztése szükséges a biztonságos üzemmenet biztosítására.

Azonban a matematikai modellek megalkotásához megfelelő adatok szükségesek, amik kiválasztása hosszadalmas és nagy szakértelmet igénylő munka. Így olyan új eszközök elkészítése szükséges, amelyek a már tárolt és éppen gyűjtött folyamatadatokat felhasználva az adatok közti kapcsolatot is figyelembe véve képes elkülöníteni a célnak megfelelő, különböző adatszegmenseket.

A dolgozat célja olyan eszközök bemutatása, amelyek közvetlenül vagy közvetve hozzájárulhatnak a termelő folyamat fejlesztéséhez a folyamatadatokat felhasználva. Egyik megközelítés, amikor a vegyipari folyamat bemeneti és kimeneti adatait, idősorait szakaszokra, szegmensekre bontjuk, egyidejűleg a köztük lévő kapcsolatot megteremtő folyamat modelljét is figyelembe véve a további analízis alapjaként. A dinamikus főkomponens-elemzés lineáris kapcsolatot teremt a bemeneti és kimeneti adatok közt. Ezt a megközelítést integrálva a klasszikus egyváltozós idősorszegmentálási technikákba egy olyan eszközt kapunk, amely alkalmas az bemeneti-kimeneti adatok közti lineáris kapcsolat megváltozásának detektálására, ami sok esetben valamilyen meghibásodásból, rendellenességből adódik. Ezen időtartományok ismerete a folyamatfejlesztés első lépcsője lehet, hiszen kiválogathatók azok az adatrészek, amelyek további matematikai modellek előrejelző képességét ronthatják. Az így kapott homogén idősorszegmenseket tovább felhasználhatjuk a technológia matematikai modelljének megkonstruálásához. Szükségessé válik azon adatok

elkülönítése is, amelyek a matematikai modell paramétereinek meghatározásakor nagy információtartalommal birnak, segítségükkel pontosan meghatározhatók ezek a bizonyos paraméterek. Ezen algoritmus során az bemeneti-kimeneti adatok közti kapcsolatot pl. a Fisher információs mátrix prezentálhatja, amely a adott bemeneti jelsorozat mellett a kimeneti adatok modellparaméterek szerinti parciális deriváltjait tartalmazza. Ennek alkalmazásával adott modellstruktúra esetén képesek lehetünk meghatározni egy-egy bemeneti adatsor információtartalmát, azaz azt az információs potenciált, amivel az adatsor a paraméterek meghatározása szempontjából rendelkezik, ezzel csökkentve az ipari kísérletek igényét a modellezési folyamat során. Emellett egy kísérlettervezéses módszeren alapuló szabályozóhangolási módszert is bemutatok, ami közvetlenül a folyamat gazdasági hatékonyságát mérve segíti a termelő vállalat nyereségnövelését erőteljes hangsúlyt fektetve a mérnöki, műszaki megközelítésre.

Abstract

Development of Experimental Design Techniques for Analyzing and Optimization of Operating Technologies

Enormous quantity of process data and implicit information are collected and stored as function of time, in sets of so-called time-series, thanks to the application of modern, computer based control systems in the chemical industry. The extraction of the hidden (not so obvious) information in the historical process data during plant-intensification and development is a relatively new field. Just a couple of years of experience is collected. The major aim of the intensification is to increase the economic benefit of the company. At the meantime the production process is keeping getting closer to its limits defined by physical and chemical laws. That is why further investigation is necessary in field of process monitoring and control to assure the safe operation. The recent fault detection, process control and optimization solutions use mathematical models for prediction. The success of these applications depends on the prediction ability of applied models. Hence it is inevitable to develop new engineering tools, which support creating accurate and robust process models. As most of the models are data driven, the proper selection of stored process data used in model construction is essential to reach success. Highlighting the goal oriented data slices is a time-series segmentation task.

The aim of this thesis is to introduce theoretical basics of different approaches which can support further the production process development, based on the extracted knowledge from process data. As selection of time-frame with a certain operation is the starting point in a further process investigation, Dynamic Principal Component Analysis (DPCA) based time-series segmentation approach is introduced in this thesis first. This new solution is resulted by integrating DPCA tools into the classical univariate time-series segmentation methodologies. It helps us to detect changes in the linear relationship of process variables, what can be caused by faults or misbehaves. This step can be the first one in the model-based process development since it is possible to neglect the operation ranges, which can ruin the prediction capability of the model. In other point of view, we can highlight problematic operation regimes and focus on finding root causes of them.

When fault-free, linear operation segments have been selected, further

segregation of data segments is needed to find data slices with high information content in terms of model parameter identification. As tools of Optimal Experiment Design (OED) are appropriate for measuring the information content of process data, the goal oriented integration of OED tools and classical time-series segmentation can handle the problem. Fisher information matrix is one of the basic tools of OED. It contains the partial derivatives of model output respect to model parameters when considering a particular input data sequence. A new, Fisher information matrix based time-series segmentation methodology has been developed to evaluate the information content of an input data slice. By using this tool, it becomes possible to select potentially the most valuable and informative time-series segments. This leads to the reduction of number of industrial experiments and their costs. In the end of the thesis a novel, economic-objective function-oriented framework is introduced for tuning model predictive controllers to be able to exploit all the control potentials and at the meantime considering the physical and chemical limits of process.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The economic crisis started in 2007 had serious effect also on the chemical industry and it is going to have effect in the next couple of years. New strategic direction has been evolved with the combination of increasing operation efficiency and high return on investment, especially in short term. The importance of energy saving, cost reduction solutions are highly appreciated and directions of these particular research fields are highly supported. The Advanced Process Control (APC), process monitoring and fault detection solutions can contribute to reach the goal either explicitly by e.g. reducing the utility consumption or explicitly by detection the occurred faults and being able to prevent further damages.

Information is a highly powerful resource to reach the previously mentioned goals. Huge amount of process data is archived thank to the highly automated chemical processes. These data archives have huge potential to extract valuable information from them for different purposes. Data collection usually takes place early in an improvement project, and is often formalized through a data collection plan, which often contains the following activities.

- Pre-collection activity (agreement on goals, target data, definitions, methods).
- Collection of data sets.
- Present Findings - usually involves some form of sorting analysis and/or presentation [1] .

In development projects of operating processes, targeted data collection usually means additional experiments to carry out. In lot of cases there is no option to make the best of extracting information from data collected in normal operation.

In Ref.[2] data collection methods and the use of historical process data in process improvement are very well described, having focus on Knowledge Discovery in Databases (KDD) framework. KDD has multiple steps to extract information from process data stored in databases, which are depicted in Figure 1.1.

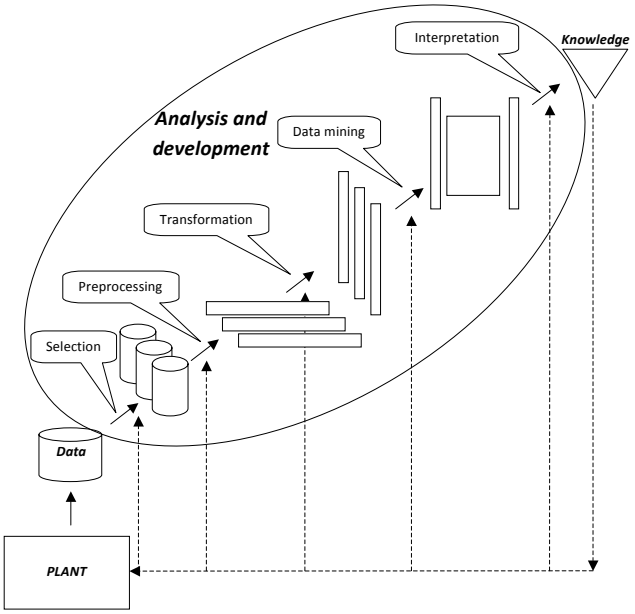

Figure 1.1: Knowledge Discovery in Databases - data driven development

System identification, fault detection and process monitoring, time series analysis are considered as engineering tasks as part of KDD. The steps of this framework can be summarized as follows:

- Data selection: similarly to the precollection activity, the main application domain is developed and the main goal of KDD process is identified.
- Data pre-processing: the step of filtering and reconciliation of the collected data, in order to correct measurement errors e.g. due to measurement noise.
- Data transformation: the step of finding and performing features and application methods on process data for achieving the set of goals. This step prepares process data for the information extraction step.
- Data mining: extraction of information from the previously transformed process data. Various methods can be applied in this step, like clustering, regression, classification or time-series segmentation. The result might be

either expected and it confirms premisses or it can open brand new directions in process development.

Online detection of any misbehavior in the technology by analyzing the recently collected process data is a well defined engineering goal (it is called knowledge in KDD). PCA is well-known and applied approach in process monitoring. Investigation of transient states needs dynamic PCA, since it can describe the dynamic behavior more accurately. Conventional PCA needs a huge amount of data for calculation of covariance matrix. Collecting enough process data to perform a new calculation of covariance matrix ruins the possibility of accurate online detection. Recursive PCA is for solving this problem since it needs only the latest process data and the latest covariance matrix to update covariance matrix. In this thesis, a combination and integration of the recursive and dynamic PCA is proposed and it is inserted into time series segmentation techniques. This results efficient multivariate time-series segmentation methods to detect locally linear homogenous operation ranges. E.g. this tool helps to reach the goal of online fault detection and separation of operation regimes, mentioned in the beginning of the paragraph. The similarity of time-series segments is evaluated based on the Krzanowski-similarity factor, which compares the hyperplanes determined by the PCA models.

Advanced chemical process engineering tools, like model predictive control or soft sensor solutions require proper process model. Parameter identification of these models needs historical data segregation of input-output data with high information content. This engineering goal is also a time-series segmentation task, like the detection of locally linear operation regimes in the previous case. Classical model based Optimal Experiment Design (OED) techniques can be applied to create experiments providing highly informative data. These solutions can also be utilized in the selection of informative segments from historical process data. In the thesis, a goal-oriented OED based (Fisher information matrix is the appropriate tool in OED) based time-series segmentation algorithm is introduced to fulfil this demand.

1.1 Time-series segmentation for process monitoring

Continuous chemical processes frequently undergo number of changes from one operating mode to another. The major aims of monitoring plant performance at process transitions are (i) the reduction of off-specification production, (ii) the identification of important process disturbances and (iii) the warning of process

malfunctions or plant faults. The first step in optimization of transitions is the intelligent analysis of archive (historical) and streaming (on-line) process data ([3]).

Fault and any misbehavior detection shall be performed by collecting and analyzing the input-output data together. This way, determination of events becomes possible where the correlation between the input and output process varies. Separation of input-output data during investigation of process events, and considering them as univariate time-series is impractical and misleads the engineer. In this case we miss considering the process itself, which should be the object of the detailed investigation.

The detailed investigation of historical and streaming process data is recognized by the process monitoring and control industry and complex solutions have been provided for reaching these goals. Principal Component Analysis (PCA) and Partial Least-Squares (PLS) ([4]) and their different modifications have been applied more and more widespread ([5, 6, 7, 8]). Process monitoring approaches based on these methods conduct statistical hypothesis tests on mainly two indices, the Hotelling T^2 and Q statistics in principal component and residual subspaces, respectively. Process failures, similarly to sensor faults ([9]) can be detected effectively by analyzing these metrics. Several approaches have been developed to extend the applicability of these methods for non-linear problems: by combining wavelet transformation and PCA tools ([10]) or by extension of PCA with kernel methods ([11]). Other multivariate process monitoring approach is the integration of qualitative trend analysis (QTA) and PCA ([12]). Correlation changes of process variables are detected by QTA, PCA assures to handle the multivariate characteristics of the process. In Ref. [13] a new methodology has been developed based on applying wavelets and combining it with Markov trees to enhance the process monitoring performance.

Nowadays nature inspired fault detection techniques become more and more widespread, like immune system or neural networks based methodologies ([14, 15, 16]). Artificial immune system based techniques mimic the operation of the human health defense system. "White blood cell" objects are defined and cloned to detect any faults in the operating technology ([15, 16]). It is similar to the principle of negative selection, where the immune system distinguish healthy body cells and foreign cells causing defections. The application of these methodologies starts with initialization where the definition of white blood cell objects is done. The next step is to choose a operation range or pattern for the learning algorithm, and to choose a test case to validate the result of the learning step. Highly experienced user has

the ability to select appropriate patterns and ranges for these purposes. Careful selection of these data sets is essential to avoid disadvantages of these methods like the constrained ability of recognition unlearnt scenarios. Artificial immune based techniques are applicable not just detecting but diagnosing the root causes of process faults. Artificial neural networks are also widely applied to establish even non-linear connection between input-output data, and to detect any misbehavior. Recent techniques from this field are summarized in details in Ref [14].

Time-series segmentation methodologies are for segregating collected process data with considering pre-defined aspect. This method is often used to extract internally homogeneous segments from a given time series to locate stable periods of time, to identify change points ([17], [18]). Although in many real-life applications a lot of variables must be simultaneously tracked and monitored, most of the time series segmentation algorithms are based on only one variable ([3]).

One of the aims in this thesis is to introduce new time series segmentation algorithms that is able to handle multivariate data sets to detect changes in the correlation structure among process variables. PCA is the most frequently applied tool to discover information in correlation structure [19] like in field of fault detection [20]. As conventional PCA model defines a linear hyperplane and most production processes are non-linear, this approach is restrictively applicable for analyzing chemical technologies. To be able to locally linearize and this way follow the non-linearity of the process, we need to use the latest data of that operation regime to calculate proper models. This model calculation can be inaccurate if considering older data points which have no information content regarding to the recent PCA model. It might be necessary since computation of PCA models needs a numerous data points. The recursive computation way of PCA models ([21]) has the capability to use only the latest data points since the recent PCA model is yielded by updating the existing PCA model with the recent measurements. By applying the variable forgetting factor developed by Fortescue et al. [22], it is possible to determine the weight of the recently collected process data points in computation of the updated and proper PCA model. As conventional PCA is for analyzing static data, Ku et al. presented a methodology [23] called dynamic PCA (dPCA) to be able to handle the time dependency of the process data. In dynamic PCA the initial data matrix has been augmented with additional columns of process data collected in previous sample times. Combining the method of recursive PCA with dynamic PCA, a helpful tool is yielded to describe the dynamic behavior of multivariate dynamic systems throughout analysis of process data. Choi and Lee

developed an algorithm with kernel functions to capture the nonlinear behavior of chemical processes and combined it with dynamic PCA to describe the dynamical characteristics ([24]).

The developed segmentation algorithm can be considered as the multivariate extension of piecewise linear approximation (PLA) of univariate data sets ([25]). Most of data mining algorithms utilize a simple distance measure to compare the segments of different time series. This distance measure is calculated based on the linear models used to describe the segments ([26]). The distance of PCA models could be determined using the PCA similarity factor developed by Krzanowski [27, 28]. By integrating the Krzanowski similarity measure into classical segmentation techniques new PCA based segmentation methods are resulted. As PCA has different metrics to describe the PCA model (Hotelling T^2 statistics, Q reconstruction error) in the further research the possibilities of utilization of these metrics are going to be investigated in details.

In the previously mentioned nature inspired methodologies ([14, 15, 16]) need learning and validating process to become ready to be used to detect and recognize process misbehaves and faults. Unlike these tools, the developed dPCA based technique needs only initialization by defining the initial variance-covariance matrix. As the developed time-series segmentation based process monitoring approach is based on continues update of PCA models and detection of differences between the PCA models, learning and validating process can be skipped. The need of PCA model update means changes in the correlation structure of input-output variables. When fault diagnosis is needed beside the fault detection, supervised learning has to be applied to learn the different fault scenarios.

The application way of the developed algorithm is presented on a simple auto-regressive (AR) system and on the benchmark of Tennessee-Eastman (TE) problem. Multivariate statistical tools are mostly evaluated and tested on Tennessee-Eastman problem in chemical and process engineering practice. TE process is applied for various goals, like [29] developed and tested his branch and bound methodology for control structure screening. In field of process monitoring several intuitive methodologies are developed and presented throughout TE example, like Bin Shams et al. [30] introduced his method based on CUSUM based PCA. The Independent Component Analysis is also widely applied for monitoring process performance, like Lee et al. [31], who also generated test faults for his methodology using the TE process. The application of one of the most intuitive methodologies is also presented using this benchmark, where Srinivasan [32] developed a fault

detection framework based on dynamic locus analysis for online fault diagnosis and state identification during process transitions.

1.2 Time-series segmentation based on information content

Most of advanced chemical process monitoring, control, and optimization algorithms rely on a process model. Some parameters of these models are not known *a priori*, so they must be estimated from process data either collected during targeted experiments or normal operation (historical process data). The accuracy of these parameters largely depends on the information content of data selected to parameter identification [33]. So there is a need for an algorithm that can support information content based selection of data sets. As we saw previously, targeted selection of suitable data sequences is a time-series segmentation task.

Shannon defined the information content first in the field of communication [34]. An essential problem of communication was to reproduce a message in the end point that was coded in the starting point. In Shannon's theory, entropy measures the information content of the message. It simply measures the uncertainty in the message just like Gibbs entropy measures the disorder in a thermodynamic system.

Information content in mathematical model identification is a specific characteristic of the chosen process data set: measures the "useful" variation of the model input data set which causes significant changes in the model output data set. "Useful" variation helps us the determine the model parameters using identification techniques. Obviously, the more "useful" variation we have, the higher information content the considered data set possesses.

Extraction of highly informative time-series segments from historical process data requires novel, goal-oriented time-series segmentation algorithm. One of the key ideas in this thesis is the propose the utilization of tools of optimal experiment design (OED) in this new field. Franceschini [35] provides an overview and critical analysis of this technique. Tools of optimal experiment design are applicable to measure the information content of datasets [36] regarding to a pre-defined process model. Usually OED is an iterative procedure that can maximize the information content of experimental data through optimization of (process model) input profiles [37]. OED is based on a sensitivity matrix - so-called Fisher information matrix, which is constituted from partial derivatives of model outputs respect to changes

of selected model parameters. Tools and methods of sensitivity analysis are summarized in details by Turanyi [38]. The information content of the selected input data sequence can be measured by utilizing A, E or D criteria [36, 39] based on the sensitivity based Fisher information matrix. As each criterion is an aggregate metric based on the Fisher information matrix, they all have their strengths and weaknesses. In order to mitigate these drawbacks and being able to combine the advantages of these metrics, Telen et al. [40] developed a multi-objective approach which enables to combine two optimization criteria. Experimental design procedures for model discrimination and for estimation of precise model parameters are usually treated as independent techniques. In order to match the objectives of both procedures, Alberton et al. [41] proposes use of experimental design criteria that are based on measures of the information gain yielded with new experiments.

Experimental design techniques are also available to design optimal discriminatory experiments [42], when several rival mathematical models are proposed for the same process. Performing additional experiment for each rival model may undermine the overall goal of optimal experimental design, which is to minimize the experimental effort. Brecht et al. deals with the design of a so-called compromise experiment [42], which is an experiment that is not optimal for each of the rival models, but sufficiently informative to improve the overall accuracy of the parameters of all rival models. Alberton et al. presents a new design criterion for discrimination of rival models [43], taking into account the number of models that are expected to be discriminated after execution of the experimental design.

In this thesis OED techniques are utilized to extract informative segments from a given time-series and separate different segments to identify different models or parameters. At first the input signal should be separated into sets of input sequences as the basis for constituting the Fisher information matrix. Fisher matrix can possess surplus information - not just the quantity of the information but its direction in the considered information space. Time series segments with similar information content can be described by similar information matrices. With the help of Krzanowski distance measure [27] it is possible to determine the similarity of Fisher matrices by direct comparison of them. In the thesis a novel and intuitive time-series segmentation algorithm is introduced for supporting to identify parameter sets from the most appropriate time frame of historical process data. This leads to reduction in the cost and time consumption of parameter estimation.

1.3 Economic based application of experiment design

The information extracted from plant historians is highly suitable for supporting controller design. Thanks to the general non-linear behavior of chemical processes of it is very difficult to find the right tuning parameters of the controllers in the whole operation range. In spite of the nonlinearity of processes most of the controllers are installed with linear algorithms. The right parameters for the production (e.g. set-points, tuning parameters of controllers, valve positions) are determined experimentally using the intuition of engineers. The response of the non-linear process is approximated with linear models and each linear model is valid only within a narrow operation range. A model library is needed to be created to characterize the operating process in the whole operation range [44]. The possibilities of the model library conducted to the demand of re-determining controller tuning parameters even if it is an iterative process like iterative learning control in batch processes [45, 46]. A know-how is necessary to fulfill these requirements, which is easy to implement even in case of operating model predictive controllers (MPCs). The tools of classical experiment design techniques can reach all these goals.

The production in the chemical industry represents a typical example of a multiproduct process. One reactor is used for producing various products and changes of production circumstances is handled by control algorithms. During transitions between products, off-specification products are produced. This product is generally worth less than the on-specification material (which fulfill all the commercial and quality requirements), therefore it is crucial to minimize its quantity. From control algorithm point of view, it means to find correct tuning parameters which enable reaching new setpoints. Beside the importance of the process transients, an on-specification product can be produced under varying circumstances and at varying operating points, which motivates us to find the (e.g. economically) optimal operation point in production stages.

The demand for time and cost reduction of grade transitions inspire researchers to find more and more innovative solutions [47, 48]. Optimization of complex operating processes generally begins with a detailed investigation of the process and its control system [31]. It is important to know how information stored in databases can support the optimization of product transition strategies. How hidden knowledge can be extracted from stored time-series, which can assure additional possibilities to reduce the amount of off-grade products. The optimization of

transition is a typical task in process industry [49]. Modern control algorithms and strategies are available to handle these tasks effectively. The determination of the tuning parameters of these algorithms are quite difficult, time and cost consuming and experimental process.

One of the common experimentation approaches is One-Variable-At-a-Time (OVAT) methodology, where one of the variables is varied while others are fixed. Such approach depends upon experience, guesswork and intuition for its success. On the contrary, tools like design of experiments (DoE) permit the investigation of the process via simultaneous changing of factors' levels using reduced number of experimental runs. Such approach plays an important role in designing and conducting experiments as well as analyzing and interpreting the data. These tools present a collection of mathematical and statistical methods that are applicable for modeling and optimization analysis in which a response or several responses of interest are influenced by various designed variables (factors) [50].

There are typical grade/ operation "sequences" during running the processes so it is possible to handle them like a "batch" in the pharmaceutical industry. This approach allows us to integrate the iterative learning control scheme into the optimization of the grade transition. It means that the optimal grade change strategy - by manipulating the tuning parameters of the controller - could not be found in one step but iteratively. The experimental design techniques need low number of iteration during optimization, so they are beneficial if combined with the iterative learning control theory.

In this thesis the applicability of experimental design technique is going to be examined. This approach will be proven to be appropriate for finding the right tuning parameters of an MPC controller. The aim of the case study is the reduction the time consumption of transitions.

Chapter 2

On-line detection of homogeneous operation ranges by dynamic principal component analysis based time-series segmentation

Any development in process technologies should be based on the analysis of process data. In the field of process monitoring the recursive Principal Component Analysis (PCA) is widely applied to detect any misbehavior of the technology. Recursive computation of PCA models means combining the existing PCA model with the recent process measurements and it results the recent PCA model. The investigation of transient states needs dynamic PCA to describe the dynamic behavior more accurately. By augmenting the original data matrix of PCA with input-output data from previous sample times conventional PCA becomes "dynamised" to catch the time-dependency of process data. The integration and combination of recursive and dynamic PCA into classic time series segmentation techniques results efficient multivariate segmentation methods to detect homogenous operation ranges based on either historical or streaming process data. By these new multivariate time-series segmentation techniques we can support process monitoring and control by separation of locally linear operation regimes. The performance of the proposed methodology is presented throughout an example of a linear process and the commonly applied Tennessee Eastman process.

2.1 dPCA based multivariate time series segmentation

In the process industry we need to be able to connect manufactured products and the operation regimes in which they were produced. It has several advantages: it becomes possible to detect occurrences of different disturbances, there is chance to find a suitable operation regime for model parameter estimation, etc. The first step on this way should be the analysis of historical process data. The correlation between input-output data is determined by the process itself, the output process data shall be handled as complex function of input process data. Hence input-output datasets can be considered as a multivariate time-series. Dynamic principal component analysis is a suitable for extending the classical univariate time-series approaches for multivariate cases and it can be the basis for developing a toolbox to fulfill demand of segregating the homogeneous operation regimes.

The chapter is organized as follows: in the rest of this section the main parts of the developed dynamic PCA based time-series segmentation methodology is going to be introduced. First, the principles of time series segmentation is explained in details. Then the connection of classical and dynamic principal component analysis and time-series segmentation is presented. In the end of the section integration of the most important components, like recursive computation of covariance matrices and Krzanowski similarity measure is shown and then the detailed summary of developed algorithm is introduced.

2.1.1 Multivariate time series segmentation algorithms

A multivariate time series $T = \{\mathbf{x}_k = [x_{1,k}, x_{2,k}, \ldots, x_{n,k}]^T | 1 \leq k \leq N\}$ is a finite set of N n-dimensional samples labelled by time points $t_1, \ldots, t_N$. A segment of T is a set of consecutive time points which contains data point between segment boarders of a and b. If a segment is denoted as $S(a, b)$, it can be formalized as: $S(a, b) = \{a \leq k \leq b\}$, and it contains data vectors of $\mathbf{x}_a, \mathbf{x}_{a+1}, \ldots, \mathbf{x}_b$. The c-segmentation of time series T is a partition of T to c non - overlapping segments $S_T^c = \{S_i(a_i, b_i) | 1 \leq i \leq c\}$, such that $a_1 = 1, b_c = N$, and $a_i = b_{i-1} + 1$. In other words, an c-segmentation splits T to c disjoint time intervals by segment boundaries $s_1 < s_2 < \ldots < s_c$, where $S_i(s_i, s_{i+1} - 1)$.

The goal of the segmentation procedure is to find internally homogeneous segments from a given time series. Data points in an internally homogeneous

segment can be characterized by a specific relationship which is different from segment to segment (e.g. different linear equation fits for each segments). To formalize this goal, a cost function $cost(S(a,b))$ is defined for describing the internal homogeneity of individual segments. Usually, this cost function $cost(S(a,b))$ is defined based on distances between actual values of time-series and the values given by a simple function (constant or linear function, or a polynomial of a higher but limited degree) fitted to data of each segment (the model of the segment). For example in [51, 52] the sum of variances of variables in segment was defined as $cost(S(a,b))$:

$$\begin{aligned} cost(S_i(a_i,b_i)) &= \frac{1}{b_i-a_i+1}\sum_{k=a_i}^{b_i} \parallel \mathbf{x}_k - \mathbf{v}_i \parallel^2, \\ \mathbf{v}_i &= \frac{1}{b_i-a_i+1}\sum_{k=a_i}^{b_i} \mathbf{x}_k, \end{aligned} \tag{2.1}$$

where $\mathbf{v}_i$ the mean of the segment.

Segmentation algorithms simultaneously determine parameters of fitted models used to approximate behavior of the system in segments, and a_i, b_i borders of the segments by minimizing the sum of costs of the individual segments:

$$cost(S_T^c) = \sum_{i=1}^{c} cost(S_i(a_i,b_i)) \,. \tag{2.2}$$

My aim in this thesis is to extend the univariate time series segmentation concept to be able to handle multivariate process data. In the simplest univariate time-series segmentation case the cost of S_i segment is the sum of the Eucledian distances of the individual data points and the mean of the segment.

In the multivariate case a covariance matrix is calculated in every sample time, so the result is a "covariance matrix time-series". The cost of S_i segment is the sum of the differences of the individual PCA models to the mean PCA model calculated from the mean covariance matrix. The similarities or differences among multivariate PCA models can be evaluated with the PCA similarity factor, Sim_{PCA}, developed by Krzanowski [27, 28]. It is used to compare multivariate time series segments. Similar to Eq(2.1), the similarity of covariance matrices in the segment to the mean covariance matrices can be expressed as the cost of the segment. Consider S_i segment with a_i and b_i borders. A covariance matrix ($\mathbf{F}_k$) is calculated in every sample point between the segment boarders, $a_i \leq k \leq b_i$ The mean covariance

matrix can be calculated as:

$$\mathbf{F}_T(a_i, b_i) = \frac{1}{b_i - a_i + 1} \sum_{k=a_i}^{b_i} \mathbf{F}_k \tag{2.3}$$

where $\mathbf{F}_k$ covariance matrix is calculated in the k^{th} time step from the historical data set having n variables. PCA models of S_i segment consist of p principal components each. The eigenvectors of $\mathbf{F}_T$ and $\mathbf{F}_k$ are denoted by $\mathbf{U}_{T,p}$ and $\mathbf{U}_{k,p}$, respectively. The Krzanowski similarity measure is used as cost of the segmentation and it is expressed as:

$$Sim_{PCA}(a_i, b_i) = \frac{1}{b_i - a_i + 1} \sum_{k=a_i}^{b_i} \frac{1}{p} \text{trace} \left(\mathbf{U}_{T,p}^T \mathbf{U}_{k,p} \mathbf{U}_{k,p}^T \mathbf{U}_{T,p} \right) \tag{2.4}$$

In the equation above (Eq(2.4)) $\mathbf{U}_{k,p}$ is calculated based on the decomposition of the $\mathbf{F}_k$ covariance matrix $\mathbf{F}_k = \mathbf{U}_k \Lambda_k \mathbf{U}_k^T$ into a Λ_k matrix which includes the eigenvalues of $\mathbf{F}_k$ in its diagonal in decreasing order, and into a $\mathbf{U}_k$ matrix which includes the eigenvectors corresponding to the eigenvalues in its columns. With the use of the first few nonzero eigenvalues ($p < n$, where n is the total number of principal components, p is the number of applied principal components) and corresponding eigenvectors, PCA model projects correlated high-dimensional data onto a hyperplane of lower dimension and represents relationship in multivariate data.

Since $\mathbf{F}_k$ represents the covariance of the multivariate process data in the k^{th} sample time, the calculation of $\mathbf{F}_k$ can be realized in different ways, e.g. in a sliding window or recursive way. In this thesis $\mathbf{F}_k$ is calculated recursively on-line, the detailed computation method is presented in Section 2.1.3.

The cost function Eq(2.2) can be minimized using dynamic programming by varying the place of segment borders, a_i and b_i (e.g. [52]). Unfortunately, it is computationally too expensive for many real data sets. Hence, usually one of the following heuristic, most common approaches are followed [25]:

- **Sliding window:** A segment is continuously growing and the recently collected data point is merged until the calculated cost in the segment exceeds a pre-defined tolerance value. For example a linear model is fitted on the observed period and the modeling error is analyzed.

- **Top-down method:** The historical time series is recursively partitioned until some stopping criteria is met.

- **Bottom-up method:** Starting from the finest possible approximation of historical data, segments are merged until some stopping criteria is met.

In data mining, bottom-up algorithm has been used extensively to support a variety of time series data mining tasks [25] for off-line analysis of process data. The algorithm begins with creating a fine approximation of the time series, and iteratively merge the lowest cost pair of segments until a stopping criteria is met. When the pair of adjacent segments $S_i(a_i, b_i)$ and $S_{i+1}(a_{i+1}, b_{i+1})$ are merged a new segment is considered $S_i(a_i, b_{i+1})$. The segmentation process continues with calculation of the cost of merging the new segment and its right neighbor and its left neighbor ($S_{i-1}(a_{i-1}, b_{i-1})$ segment) and then with further segment merging.

To develop a multivariate time-series segmentation algorithm which is able to handle streaming process data, sliding window approach should be followed. After initialization, the algorithm merges recently collected process data to the existing segments until the stopping criterion is met. The stopping criterion is usually a determined value of the maximal merging cost.

This algorithm is quite powerful since merging cost evaluations requires simple identifications of PCA models which is easy to implement and computationally cheap to calculate. The sliding window method is not able to divide up a sequence into a predefined a number of segments; on the other hand this is the fastest time-series segmentation method.

2.1.2 Application of PCA for analyzing dynamic systems

The classical PCA is mainly for exploring correlations in data sets without any time dependency. In some industrial segments (e.g. in some polymerization processes) time consumption of process transitions is in the same order of magnitude with the length of a steady state operation. Hence it is crucial to be able to analyze and extract information from data sets collected in transitions. The demand of being able to handle time dependency of the collected process data motivated Ku and his colleagues [23] to dynamize the static PCA for the needs of dynamic processes. Consider the following process:

$$\mathbf{y}_{k+1} = \mathbf{a}_1\mathbf{y}_k + \ldots + \mathbf{a}_{n_a}\mathbf{y}_{k-n_a} + \mathbf{b}_1\mathbf{u}_k + \ldots + \mathbf{b}_{n_b}\mathbf{u}_{k-n_b} + \mathbf{c} \qquad (2.5)$$

where $\mathbf{a}_i$, $\mathbf{b}_j$ ($i = 1, \ldots, n_a, j = 1, \ldots, n_b$) and $\mathbf{c}$ are vectors of constants, n_a and n_b show the time dependency of process data, $\mathbf{u}_k$ is the vector of k^{th} sample of

(multivariate) input and $\mathbf{y}_k$ is the output (product) vector in the same time. Ku et al. [23] pointed out that performing PCA on the $\mathbf{X} = [\mathbf{y}, \mathbf{u}]$ data matrix preserves the auto and cross correlations caused by time variance of the time series such as the ones above.Thus, Ku et al. [23] suggested that the $\mathbf{X}$ data matrix should be formed by considering the process dynamics at every sample point. Generally speaking, every sample point should be completed with the points they are depending on, i.e. the past values:

$$\mathbf{X} = \begin{bmatrix} \mathbf{y}_k & \dots & \mathbf{y}_{k-n_a} & \mathbf{u}_k & \dots & \mathbf{u}_{k-n_b} \\ \mathbf{y}_{k+1} & \dots & \mathbf{y}_{k-n_a+1} & \mathbf{u}_{k+1} & \dots & \mathbf{u}_{k-n_b+1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_{k+m} & \dots & \mathbf{y}_{k-n_a+m} & \mathbf{u}_{k+m} & \dots & \mathbf{u}_{k-n_b+m} \end{bmatrix} \tag{2.6}$$

Process dynamics create relationship between inputs and outputs - relations are preserved under PCA - and a model can be fitted with a model certain order. Usually $n_a + n_b$ is higher than the model order that is fitted to the data set. $n_a + n_b$ is equal to the total number of principal components, n. Performing PCA on the modified data matrix moves unwanted correlations to noise subspace. Possible combinations of time dependence are presented in the data matrix and we select the most important combinations of these by using PCA. First zero (or close to zero) eigenvalue shows linear relationship between variables revealed by the eigenvector belongs to this eigenvalue. The method of dynamizing the PCA is recognized and effectively applied in the field of process model identification ([53, 54]) so it proves the relevance and applicability of dPCA in handling even streaming multivariate process data.

2.1.3 Recursive PCA with variable forgetting factor

To detect changes in the process dynamics in time, we have to reach the acceptable resolution which needs us to compute a PCA model in every sample time. This is also necessary in sliding window segmentation technique or in the fine approximation of bottom-up approach. To reach this goal the method of recursive PCA is applied ([55, 21]). This is based on recursively updating the variance-covariance matrix ($\mathbf{X}_k^T\mathbf{X}_k$), where $\mathbf{X}_k$ is a data matrix (like Eq (2.6)) in the k^{th} time step comprising p variables and n samples, proposed by Li et al. [21]. Recursive calculation of the (d)PCA model help us to avoid the excessive expansion of data matrix caused by frequently collected process data. This is a well-known issue in

e.g. adaptive control [22].

The proposed algorithm for recursive computation of covariance matrices is as follows:

1. Initialization

   Nominal data, $\mathbf{X}$, is defined, as previously introduced. The data set is normalized to zero mean and unit standard deviation, $\mathbf{X}_0$. Vectors of mean values, $\overline{\mathbf{x}}_0$, and standard deviation, $\mathbf{s}_0$, are saved. The initial variance-covariance matrix can be expressed as:

   $$F_0 = \frac{\mathbf{X}_0^T \mathbf{X}_0}{l-1} \tag{2.7}$$

   where l is the number of initial samples.

   The initial forgetting factor is defined as:

   $$\lambda_0 = 1 - (\frac{1}{l}) \tag{2.8}$$

2. Application of variable forgetting factor in calculation of new covariance matrices.

   (a) The new vector of measurements, $\mathbf{x}_k$, is collected. The new mean vector is calculated as:

   $$\overline{\mathbf{x}}_k = \lambda_{k-1}\overline{\mathbf{x}}_{k-1} + (1-\lambda_{k-1})\mathbf{x}_k \tag{2.9}$$

   (b) The difference of the new mean, $\overline{\mathbf{x}}_k$, and the old mean, $\overline{\mathbf{x}}_{k-1}$, is stored in $\delta\overline{\mathbf{x}}_k$. It is necessary since the recursive calculation of process variance (σ) is as follows:

   $$\sigma_k = \lambda_{k-1}(\sigma_{k-1} + \delta\overline{\mathbf{x}}_k^2) + (1-\lambda_{k-1})(\mathbf{x}_k - \overline{\mathbf{x}}_k)^2 \tag{2.10}$$

   The standard deviation in sample time k can easily be calculated from the process variance.

   (c) Next step is to normalize recently collected process data, $\mathbf{x}_k$, using previously calculated mean, $\overline{\mathbf{x}}_k$, and standard deviation, $\mathbf{s}_k$. The recursive calculation of variance-covariance matrix is formulated as:

$$F_k = \lambda_{k-1} F_{k-1} + (1 - \lambda_{k-1})(\chi_k^T \chi_k) \tag{2.11}$$

where χ_k is the normalized process data vector, $\chi_k = \frac{\mathbf{x}_k - \overline{\mathbf{x}}_k}{\mathbf{s}_k}$.

(d) The final step in the recursive calculation loop is to calculate the value of the variable forgetting factor, λ_k. To calculate the value of the forgetting factor Fortescue [22] proposed an algorithm, it is calculated as follows:

$$\lambda_k = 1 - \frac{[1 - \frac{T_k^2}{p}]\frac{e_k^2}{p}}{\sqrt{n_{k-1}}} \tag{2.12}$$

where p is the number of applied principal components, n_k is the asymptotic memory length at $(k-1)$, the T^2 is the Hotelling T^2 metric at sample point k and the error term e_k is the Q metric, the reconstruction error at sample point k presented in [55].

As the forgetting factor decreases, the recent observation get more weight in calculation of updated variance-covariance matrix with less weight being placed on older data. Hence it is one of the basic component for quick adaptation of dPCA models to describe the correlation structure in the changed operation regime. It can be handled as an indicator where dPCA model needs to be updated rapidly to describe the new relationship of process variables.

2.1.4 Recursive dPCA based time-series segmentation

The dynamic principle component analysis was introduced in the previous section as an approach to be able to handle the time dependence of the collected process data (Eq (2.6)). Applying the recursive calculation method (Eq (2.11)) a new dPCA model becomes accessible in each sample point. With the application of the variable forgetting factor (Eq (2.12)) it becomes possible to exclude as much information as included by the recent measurements.

The next step is to find a valid dPCA model for each segment - so-called mean model (Eq(2.3))- and compare the recently computed dPCA models to the mean model. The comparison of dPCA models represented by the variance-covariance matrices become possible by using the Krzanowski similarity measure (Eq (2.4)). Application of segmentation algorithms become available by the help of this similarity measure so thus the segments with different dynamic behavior can be

differentiated.

For off-line application the bottom-up segmentation method is applied. The pseudocode for algorithm is shown in Algorithm 2.1.

Algorithm 2.1 Bottom-up segmentation algorithm

0: Calculate the covariance matrices recursively and split them into initial segments (define initial a_i and b_i segment boundary indices).
0: Calculate the mean model of in the initial segments (Eq(2.3)).
0: Calculate the cost of merging for each pair of segments:
$mergecost(i) = Sim(a_i, b_{i+1})$
while actual number of segments $>$ desired number of segments **do**
Find the cheapest pair to merge:
$i = argmin_i(mergecost(i))$
Merge the two segments, update the a_i, b_i boundary indices
Calculate the mean model of in the new segment (Eq(2.3)).
recalculate the merge costs.
$mergecost(i) = Sim_{PCA}(a_i, b_{i+1})$
$mergecost(i-1) = Sim_{PCA}(a_{i-1}, b_i)$ where Sim_{PCA} is the Krzanowski distance measure
end while

The previously introduced bottom-up segmentation technique is applied as off-line time-series segmentation procedure. There are some difficulties during the application of this methodology like the determination of initial and desired number of segments. Stopping criterion of segmentation procedure can be either the desired number of segments (as introduced in Algorithm 2.1) or reaching the value of a pre defined maximal cost. If the number of the desired segments are lower than the number of different operation regimes in the considered time scale, the result of the segmentation procedure might be misleading, since two or more similar, adjacent operation regime segments can be merged. If the number of the desired segments is too high, there will be the possibility to create false segments. False segments are subsegments of a homogenous segment and are not going to be merged. The introduced dPCA based bottom-up segmentation algorithm can handle this problem, since it is convergent. It means to reduce the possibility of false segments by "collecting" borders of the false segments next to the border of the homogeneous operation regime. In details: Assume that a process transient causes changes in the correlation structure of input-output variables. So we are getting from one operation range to an other. When the process is adapting to new operation conditions the dPCA models are continuously updated. Thanks to the variable forgetting factor the speed of this adaptation is "fast". Similarities of continuously

computed dPCA models to the average dPCA model of a homogeneous segment are low during process adaptation. It is because, correlation of input and out variables is continuously changing in transient state until it gets to the new homogeneous operation range. Hence merge costs are the highest in the transient time stamps. As transient state typically cannot be described by a linear PCA model, every PCA model is significantly different from each other as well as different from PCA model in homogeneous operation. It is the cause of the convergence. Taking the value of forgetting factor into consideration in segmentation algorithms, remaining superfluous and misleading segment boarders can be distinguished. If the value of forgetting factor is rapidly decreasing and exceeds a certain limit, the boundary of the segment could be considered as a valid segment boarder, otherwise it might be considered as a false segment boarder. The number of initial segments is up to definition but it can be stated that finer approximation of the time series result more sophisticated result. Too fine approximation might ruin the robustness of the algorithm. The only constraint is the number of data points, which have the ability of defining the model of the initial segments. In this particular segmentation methodology it is possible to define one variance-covariance matrix as an initial segment.

For on-line application the sliding window segmentation method is suitable. The pseudocode of developed algorithm for multivariate streaming data is shown in Algorithm 2.2.

Algorithm 2.2 Sliding window segmentation algorithm

```
0: Initialize the first covariance matrix.
   while not finished segmenting time series do
      Collect the recent process data.
      Calculate recent the covariance matrix recursively.
      Determine the merge cost (Sim_PCA) using the Krzanowski measure.
      if S < maxerror then
         Merge the collected data point to the segment.
         Calculate the mean model of in the segment (Eq(2.3)).
      else
         Start a new segment.
      end if
   end while
```

The possible differences in the results of the off-line and on-line algorithms is caused by the totally different operation methodology, since these approaches are heuristic in terms of minimizing the cost function in a segment. Thanks to the

heuristic approach certain parameters of the algorithms are needed to be defined (e.g. the number of segments in off-line case and the pre-defined error in case of on-line approach), which might also lead to different conclusions. In general, the results are quite similar, the possible differences make us investigate the roots of the small variance in them.

2.1.5 Application of confidence limits in dPCA based process monitoring

Confidence limits for Q reconstruction error and Hotelling T^2 statistics are usually defined in the commonly applied PCA based process monitoring techniques. Augmenting the dPCA based time-series segmentation methodology with the utilization of the confidence limits lead to a complex process monitoring tool. It enables further and more investigation of the segmentation results. The confidence limits can be calculated recursively, similarly to the covariance matrix.

For Hotelling T^2 statistics the confidence limit is defined as follows:

$$CL_{T^2} = \frac{(r-1)^2}{r} \cdot B_{\alpha,\frac{p}{2},\frac{r-p-1}{2}} \tag{2.13}$$

where r is the number of already collected and examined samples ($r = k \dots m$, see Eq (2.6)), p is the number of principal components, α is the probability of false alarm for each point plotted on the control chart $B_{\alpha,\frac{p}{2},\frac{r-p-1}{2}}$ is the $(1-\alpha)$ percentile of beta distribution with parameters u_1 and u_2 ([56, 57]).

For Q reconstruction error a similar limit can be defined with the following expression:

$$CL_Q = \theta_1 \left[\frac{\eta_\alpha \sqrt{\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}} \tag{2.14}$$

where

$$h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2} \tag{2.15}$$

and

$$\theta_d = \sum_{i=p+1}^{n} \gamma_i^d = trace(R_k^d) - \sum_{j=1}^{p_k} \gamma_j^d \tag{2.16}$$

where η_α is the normal deviation corresponding to the upper $(1-\alpha)$ percentile,

n is the number of variables, R_k is the covariance matrix in the k^{th} sample time, γ_i is the eigenvalue of the i^{th} principal component.

As a covariance matrix is defined in every sample time, Hotelling T^2 and Q prediction error could be applied as indicators of the process in every sample time. Hotelling T^2 represents the movement of the data in the multidimensional space, it contains important information about the process although the variables from which it is calculated is not independent. Utilizing the Krzanowski similarity factor to compare the defined hyperspaces, the homogeneous operation segments can be segregated.

2.2 Case studies

Each component of the proposed dPCA based time-series segmentation is investigated earlier, like the recursive computation of the covariance matrix ([21]) and variable forgetting factor, defined by [22]. The novelty in the proposed methodology is the new way of application and integration of the well-known methods.

The use of the proposed time series segmentation methodology will be demonstrated throughout a simple, multivariate process and as a second a much more complex and realistic Tennessee Eastman process. Data preprocessing methods are not used in these case studies as having synthetical data sets.

2.2.1 Multivariate AR process

Problem formulation

Consider the following process, as a benchmark of Ku et al. [23]:

$$\mathbf{z}_k = \begin{pmatrix} 0.118 & -0.191 \\ 0.847 & 0.264 \end{pmatrix} \mathbf{z}_{k-1} + \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} \mathbf{u}_{k-1}, \tag{2.17}$$

$$\mathbf{y}_k = \mathbf{z}_k + \mathbf{v}_k \tag{2.18}$$

where $\mathbf{u}$ is the correlated input:

$$\mathbf{u}_k = \begin{pmatrix} 0.811 & -0.226 \\ 0.477 & 0.415 \end{pmatrix} \mathbf{u}_{k-1} + \begin{pmatrix} 0.193 & 0.689 \\ -0.320 & -0.749 \end{pmatrix} \mathbf{w}_{k-1}, \tag{2.19}$$

The process description is formulated in the very same way as Ku et al. [23] presented. The input $\mathbf{w}$ vector is a random noise with zero mean and variance 1. The output vector, $\mathbf{y}$ is equal to $\mathbf{z}$ vector and an added random noise, $\mathbf{v}$ with 0 mean and variance 0.1. The values of $\mathbf{y}$ and $\mathbf{u}$ vectors are collected as process variables. The data matrix for dynamic PCA is constructed as $[\mathbf{y}_k^T\ \mathbf{y}_{k-1}^T\ \mathbf{u}_k^T\ \mathbf{u}_{k-1}^T]$. Based on Ku's examinations ([23]) five principal component are applied, since the fourth and fifth scores still show certain auto- and cross-correlation. The remaining three scores are independent form each other.

1000 samples from normal operation data are applied for the analysis and the first 100 samples are utilized to compute the initial covariance matrix. The following scenarios are considered in the examined time scale:

1. at the 400^{th} sample time: parameters of A matrix (coefficient matrix in Eq(2.17) has been changed to
$$\begin{pmatrix} 0.380 & -0.250 \\ 0.147 & 0.264 \end{pmatrix}$$

2. at the 600^{th} sample time: means of $\mathbf{w}$ are changed from the mean of $\mathbf{w}$=0 to mean of w_1=1 and the mean of w_2=-1

3. at the 800^{th} sample time: parameters of A matrix has been changed to
$$\begin{pmatrix} 0.500 & -0.500 \\ 0.200 & 0.264 \end{pmatrix}$$

So the correlation structure of input-output data changes at the 400^{th} and 800^{th} sample. This leads to the expectations that 3 different operation segment shall be found during the segmentation procedure.

Results of the time-series segmentation

Both of the different time-series segmentation methodologies have been applied during examinations: as first, the off-line bottom-up technique (Algorithm 2.1) and then the sliding window segmentation technique for on-line purposes (Algorithm 2.2). As in traditional process monitoring the Hotelling T^2 and Q reconstruction error metrics are widely applied, so the values of these indicators also examined to compare traditional result to the proposed approach. For calculating confidence limits Eq (2.13-2.14) are utilized.

The investigated multivariate time-series is depicted in Figure 2.1.

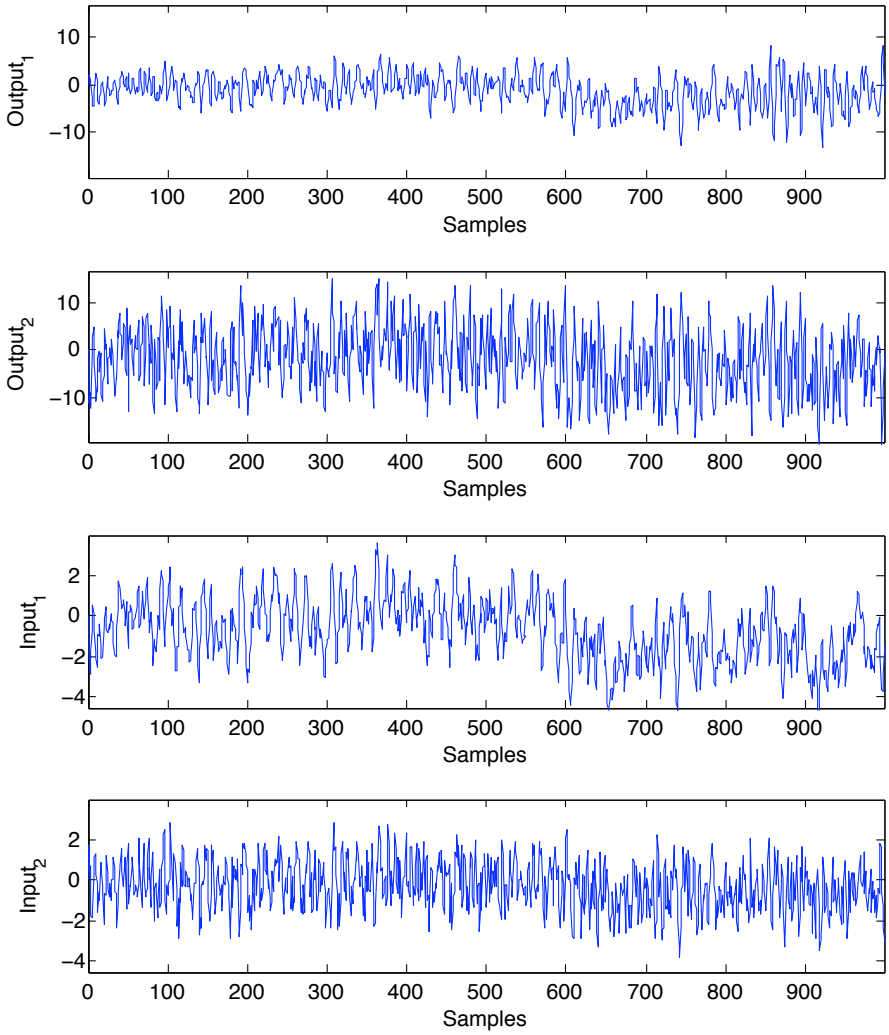

Figure 2.1: Process data in considered scenario of the AR process

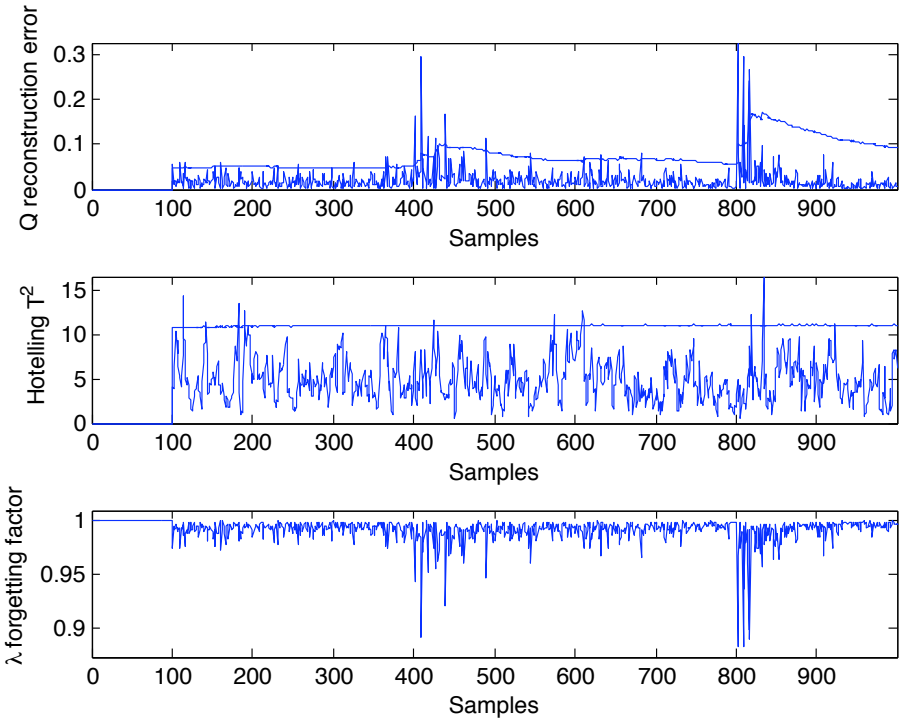

Figure 2.2: Hotelling T^2, Q metrics and value of forgetting factor in the considered time scale of the AR process

As the first 100 samples were chosen to initialize the covariance matrix, the value of these indicators is 0 in these sample times as depicted in Figure 2.2.

Changes in correlation structure are clearly detectable as it has been depicted in Figure 2.2. For quick adaptation of the dPCA model the value of the forgetting factor should be decreased in these sample times as it can be seen in Figure 2.2. The introduced algorithm detects these changes in the correlation structure so the value of the forgetting factor is automatically decreased.

In the proposed scenario, there are two changes in the correlation structure: the first at the 400^{th} sample, the second at the 800^{th} sample. The mean change of $\mathbf{w}$ at 600^{th} sample is not considered as a major difference in correlation structure as just the bias of models are different. That is why three segments are expected. If the desired number of segments are low (4 desired segments in the first case), boarders of segments are clearly identified, however segment boarders are a little shifted compared to the appearance of disturbance. It is depicted in Figure 2.3. The quantity of the delay is approximately equal to the time constant of the system. In the representation of segmentation results any kind of process data might be

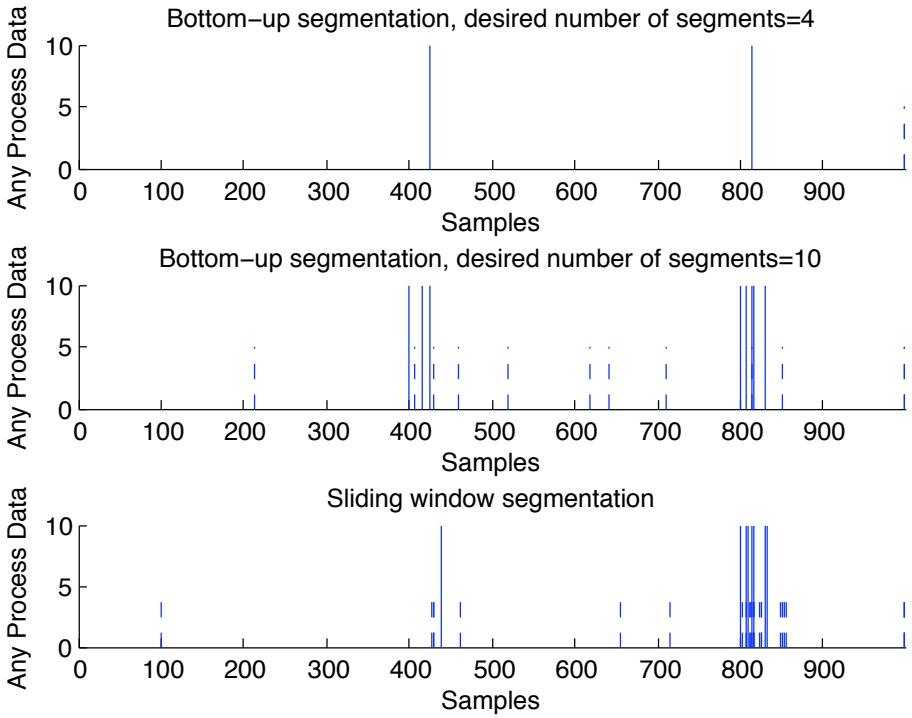

Figure 2.3: Results of different segmentation scenarios of AR process

substituted to the y axis with the proper minimal and maximal value of scale. To visualize results, y axis is rescaled with arbitrary minimal and maximal values (0 and 10, respectively).

If the desired number of segments is much higher than the number of different operation regimes in the examined time scale (like 10 in this case), it will be necessary to check to possible the extra-segment detection using the previously proposed approach. It can be handled with checking the value of forgetting factor at borders of segments. If the value of the forgetting factor does not vary significantly, the segment border can be considered as a false detection. It is shown with dashed line in Figure 2.3. A constraint for false border detection can be defined as $C = mean(\lambda) - 3\sigma$, where σ is the standard deviation of λ (in this particular case it is 0.96).

So-called quasi-segments could be detected when homogeneous operation segments are segregated. In these quasi-segments the describing dPCA model is permanently changing e.g. the system is in a transient state because of a disturbance or changing the operation point. In these cases the value of the forgetting factor decreases to assure and indicate the quick adaptation of dPCA model and crosses

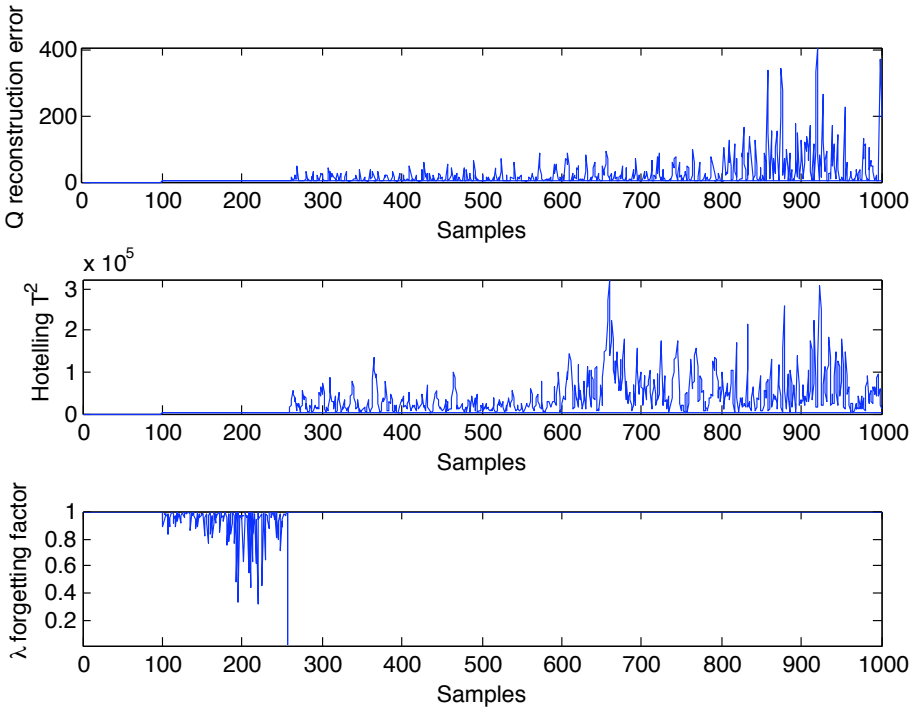

Figure 2.4: Q reconstruction error and Hotelling T^2 metrics and value of forgetting factor in the considered time scale using static PCA

the pre-defined limit. This occurs between $400^{th} - 426^{th}$ samples and $800^{th} - 831^{th}$, where dPCA model needs to adapt to changes in the correlation structure, in the $\mathbf{A}$ matrix.

As the next step the applicability of sliding window time-series segmentation methodology ($Algorithm$ 2.2) is investigated. By utilizing this approach, we got similar result as in off-line segmentation scenario, it is depicted in Figure 2.3. Similarly to the bottom-up segmentation case, the quasi-segments also could be detected, like a segment between $800^{th} - 832^{th}$.

The same examinations and segmentation scenarios are carried out with the conventional static PCA to confirm benefits of using dPCA. In this case the data matrix is constituted in the $[\mathbf{y}_k^T \ \mathbf{u}_k^T]$ form. In our examination the first 3 principal components found to have the largest eigenvalues, explained variance of 97 %As first the conventional process monitoring metrics (Hotelling T^2 and Q reconstruction error) are registered to detect the changes in the correlation structure. The result is depicted in Figure 2.4.

It is not possible to detect the changes in the correlation structure using static PCA . In this case the value of forgetting factor does not predict to find the place

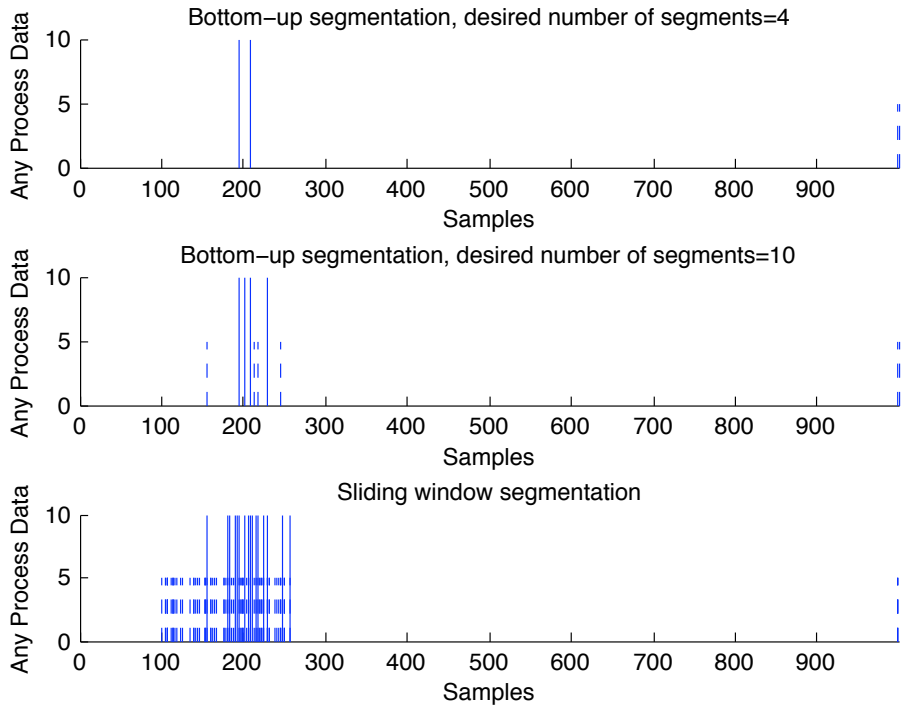

Figure 2.5: Results of different segmentation scenarios of AR process using static PCA

of changes in correlation structure, since in the 400^{th} and 800^{th} sample time the value of the forgetting factor is 1. The result of time-series segmentation scenarios strengths the previous conclusion as it is depicted in Figure 2.5. However the segment boarders are convergent to each other, but their place is not even close to the real place of changes in the correlation structure. As a conclusion the necessity of dPCA is stated.

2.2.2 The Tennessee Eastmen process

The Tennessee Eastmen problem is widely applied in the chemical and process engineering practice to test the industrial applicability of the developed process monitoring techniques([8, 29, 30, 32]). This benchmark problem has all the characteristics that an operating chemical process does, so it is suitable to evaluate the performance of the developed time series segmentation methodology.

The Tennessee Eastmen process consist of five major unit operations: a reactor, a product condenser, a vapor-liquid separator, a recycle compressor and a product gas stripper. Two products are produced by two simultaneous gas-liquid exothermic

reactions and a byproduct is generated by two additional exothermic reactions. The process has 12 manipulated variables, 22 continuous measurements and 19 composition measurements. The process is sampled with the sample time of $0.1h$. The simulator of the process was developed by Downs and Vogel in [58]. The control system used for dynamic simulations is decentralized control strategy created by Ricker [59]. The simulator includes a set of programmed disturbances listed in Table A.1 in Appendices. To be able to utilize the dynamic PCA the process variables listed in Table A.2.

To check the performance of the proposed time-series segmentation methodologies (in Algorithm 2.1 and Algorithm 2.2) the following operation scenario is considered with disturbances included: step in A/C feed ratio at 40^{th} hour, random variation in C feed temperature at 60^{th} hour, slow drift in reaction kinetics at 80^{th} hour, sticking of condenser cooling water valve at 100^{th} hour and an unknown type disturbance at 120^{th} hour ($1^{st}, 10^{th}, 13^{th}, 15^{th}$ and 19^{th} disturbance in Table A.1 in Appendices). This way 6 different segments are expected if the disturbances change the correlation structure of the input-output variables.

Data matrix is constructed as $[\mathbf{y}_k^T\ \mathbf{y}_{k-1}^T\ \mathbf{u}_k^T\ \mathbf{u}_{k-1}^T]$ to build the dPCA model . Throughout our segmentation process the first 31 principal components were applied which explain 97% of process variance. 15000 samples (150 hours long) from normal operation data are applied for the analysis and the first 3000 samples are utilized to compute the initial covariance matrix.

Results of the time-series segmentation

Similarly to the previous case, conventional process monitoring indicators - the Hotelling T^2 and Q reconstruction error - are applied to detect the disturbance introduced above. Since the first 3000 (30 hours long time scale) samples are applied to initialize the covariance matrix, the value of these indicators is 0 in these sample times as depicted in Figure 2.6.

The confidence limits for the process are determined by using Eq (2.13) and Eq (2.14). Changes in the correlation structure are detectable either in the Hotelling T^2 metric or the Q reconstruction error. These metrics cannot be utilized apart from each other: e.g. the random variation in C feed temperature is not detected in Hotelling T^2 plot but in the Q reconstruction error plot and e.g. slow drift in reaction kinetics cannot be detected just by observing Q reconstruction error plot since it can be detected in Hotelling T^2 plot. To follow the adaptation of the dPCA

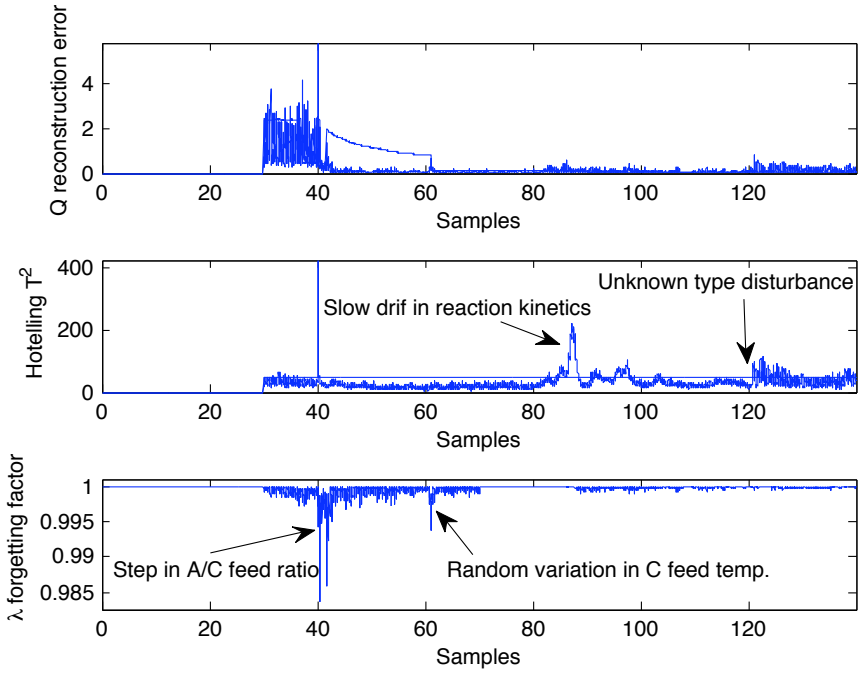

Figure 2.6: Hotelling T^2, Q metrics and value of forgetting factor in the considered time scale of TE process

model to the changes in the correlation structure the value of forgetting factor is examined, depicted in Figure 2.6.

As the forgetting factor shows having the most significant necessity for adaptation in the 40^{th} and 60^{th} hour when the step change in A/C feed ratio and random variation in C feed temperature occur. In the rest of the considered scenario the value of the forgetting factor does not change relevantly (it is close to 1). It indicates the correlation structure does not change as significantly as in the previously mentioned cases.

At first the off-line segmentation is evaluated.In the considered scenario the number of expected segments is six assuming that every disturbances change the correlation structure. This expectation should be modified after the examination of Figure 2.6, which hint less than six different operation regimes. The most important question is to determine the number of desired segments. Two cases were examined: in the first case the number of desired segments is 10, in the second one it is 20.

The results of the bottom-up scenarios are summarized in Figure 2.7. It confirms the expectations based on the examination of Figure 2.6, which means that the most significant change in the correlation structure occurs in the 40^{th} hour. Effect of

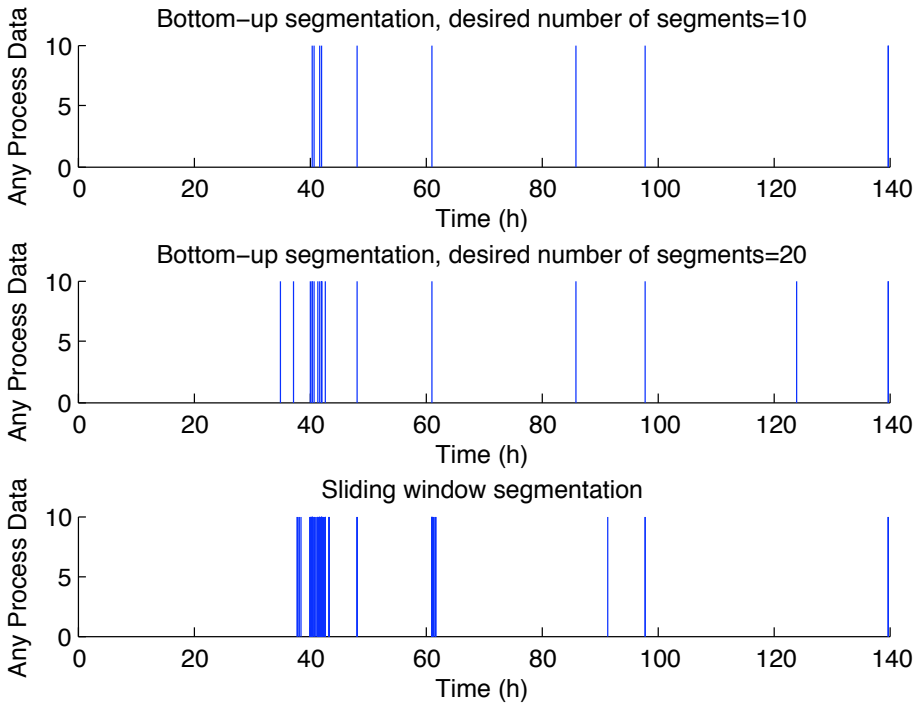

Figure 2.7: Results of different segmentation scenarios of TE process

rest of disturbances is significantly lower, however they can be detected even with using low number assumed segments. The only exception is the unknown type disturbance, occurs in the 120^{th} hour, which can not be detected in case of 10 desired number of segments. It is in accordance with small changes of the forgetting factor in that particular time scale, depicted in Figure 2.6. The number of false segment border detection is quite low despite of defining high number of desired segments.

As second the on-line segmentation methodology is applied (Algorithm 2.2). The result of the time-series segmentation scenario is shown in Figure 2.7. It is quite similar to the bottom-up scenario with 10 desired segments, since just the unknown type disturbance is not detected and every other disturbances are indicated. The closely adjacent segment borders (like around 40^{th} hour) indicates that in that particular time scale the considered system is in a transient state and adapts to the occurred disturbance.

Both of the off-line and on-line methodologies are capable to detect changes in correlation structure and rate their effects. The change of step in A/C feed ratio at 40^{th} hour and random variation in C feed temperature at 60^{th} hour can be easily detected as depicted in Figure 2.6 in the plot of the forgetting

factor. These disturbances cause significant changes in correlation structure. The rest of the disturbances cannot cause such change in the correlation structure (since the forgetting factor does not decrease as significantly as before), but e.g. detailed investigation of the classical of Hotelling T^2 metric we can highlight these disturbances. The off-line time-series segmentation algorithm is convergent, which means that the most significant disturbances are detected first and the less significant in the end. It is shown in Figure 2.7, in the first and second plot. The segment borders converged to the most significant disturbance as first (step in A/C feed ratio at 40^{th}) and then detect the other disturbances. This convergence is the reason of multiple ("thicker") segment borders. The unknown type disturbance at 120^{th} hour has the less significant effect on the correlation structure, hence it is not detected in case when the number of desired segments was 10.

The same statements can be taken by considering the on-line version of the proposed methodology. Besides detecting the changes in the correlation structure, this algorithm is also capable to rank the effect of the disturbances, however it detects the effect of the slow drift in reaction kinetics (occurred in 80^{th} hour) later than the off-line version, depicted in Figure 2.7, in the third plot. Similarly to the off-line version it neglects the effect of the unknown type disturbance at 120^{th} hour. These capabilities make the on-line version applicable for real time detection and ranking the effect of occurred disturbances, which can support the effective operation of the chemical technologies.

2.3 Summary of dPCA based time-series segmentation

In this chapter a Dynamic Principal Component Analysis based time-series segmentation framework is introduced for detecting changes of correlation structure in multivariate historical process data. Changes in the correlation structure can be caused by e.g. different kind of disturbances or faults. The basis idea of our approach is that PCA is one of the most frequently applied tool to discover information in correlation structure [19] like in field of fault detection [20]. Conventional computation of PCA models needs a numerous data points which leads inappropriate resolution of the time-series and inaccurate detection of possible faults. The recursive computation way of PCA models ([21]) has the capability to use only the latest data points since the recent PCA model is yielded by updating

the existing PCA model with the recent measurements. It results high resolution of PCA models for more accurate detection in changes of correlation structure of input-output variables. The utilization of variable forgetting factor - developed by Fortescue et al. [22] - makes possible to determine the weight of the recently collected process data points in computation of the updated and proper PCA model. As conventional PCA is for analyzing static data, we used the so-called called dynamic PCA approach presented by Ku et al. [23] to handle the time dependency of the process data.

By the integration all of the tools mentioned above we developed novel time-series segmentation algorithms for detecting changes in the correlation structure of multivariate time-series. We proposed an on-line approach for streaming data and off-line application for historical process data. In the segmentation framework the Krzanowski measure is utilized, as cost function.

For on-line applications the sliding window segmentation method is applied, for off-line applications the bottom-up technique is chosen. A simple autoregressive process and the more realistic Tennessee Eastmen process with pre-defined disturbances are used to generate input-output data sets for time-series segmentation in the case studies. Both of the off-line and on-line methods successfully detected the changes in the correlation structure with low number of false detections. The algorithm is effectively capable to detect and rank the occurred disturbances, and with the help of segregating different operation ranges discovered knowledge can be used to develop various, even cost effective operational strategies.

Chapter 3

Fisher information matrix based segmentation of multivariate data for supporting model identification

Advanced chemical process engineering tools, like model predictive control or soft sensor solutions require proper, goal oriented process models. The applicability of the model depends on its prediction ability. During parameter estimation, we determine model parameters to be able to predict the future behavior of the considered process. Parameter identification of these models needs input-output process data with high information content. Traditionally model based optimal experimental design techniques are applied to generate these data sets. When the traditional way cannot be followed, the extraction of informative segments from historical data can also support system identification. In this chapter a newly developed, goal-oriented Fisher information based time-series segmentation algorithm has been described, aimed at selecting informative segments from historical process data. The Fisher information matrix is inserted into standard bottom-up time-series segmentation approach to augment its capabilities of handling multivariate data sets and consider the correlation between them predefined by a mathematical model. Different segments can support the identification of parameter sets in different order of magnitude. Hence, we propose the use of Krzanowski's similarity coefficient between the eigenvectors of the Fisher information matrices obtained from the sequences and using either D- or E-optimality as the criterion for comparing the information content of two input sequences (neighboring segments). The efficiency of the proposed methodology

is demonstrated by two application examples. The algorithm is capable to extract segments with parameter-set specific information content from historical process data.

The chapter is organized as follows. In Section 3.1, tools of the developed algorithms are presented, like tools of Optimal Experiment Design (OED), the calculation ways of sensitivities, then the developed algorithm is described in details. In Section 3.2 the applicability of the proposed time-series segmentation method is presented through demonstrative examples.

3.1 Optimal Experiment Design based time series segmentation

Time-series segmentation extended with Optimal Experiment Design (OED) tools is highly suitable to extract data sequences with high information content.

Information content in mathematical model identification is a specific characteristic of the chosen process data set: measures the "useful" variation of the model input data set which causes significant changes in the model output data set. "Useful" variation helps us the determine the model parameters using identification techniques. Obviously, the more "useful" variation we have, the higher information content the considered data set possesses.

Since extracted data is used for parameter identification its information content should be measured. It is based on Fisher information matrix which is constructed from the sensitivities of the model output respect to the parameters. Various ways of calculating sensitivities are presented. At the end of the section the resulted time-series segmentation algorithm is described in details.

3.1.1 Background of model based optimal experimental design

The Fisher information matrix ($\mathbf{F}$) is based on the sensitivity of the model output ($\mathbf{y}(\mathbf{u}(t))$ respect to parameters ($\mathbf{p}$), calculated as follows:

$$\mathbf{F} = \frac{1}{t_{end}} \int_{t=0}^{t_{end}} \frac{\partial \mathbf{y}}{\partial \mathbf{p}}\Big|_{p=p^0}^{T}(t) \cdot \frac{\partial \mathbf{y}}{\partial \mathbf{p}}\Big|_{p=p^0}(t)dt \tag{3.1}$$

The calculation of derivatives require a process model:

$$\frac{d\mathbf{x}(t)}{dt} = f(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \tag{3.2}$$

$$\mathbf{y}(t) = g(\mathbf{x}(t)) \tag{3.3}$$

where $\mathbf{u} = [u_1 \dots u_{nu}]$ (n_u is the number of inputs) is the vector of manipulated inputs, $\mathbf{y} = [y_1 \dots y_{ny}]$ (n_y is the number of outputs) is the vector of outputs, $\mathbf{x} = [x_1 \dots x_{nx}]$ (n_x is number of states) represents the states, and $\mathbf{p} = [p_1 \dots p_{np}]$ (n_p is the number of inputs) denotes the model parameters.

Parameters in $\mathbf{p}$ are unknown and should be estimated using the data taken from experiments. The estimation of these parameters is based on the minimization of the square error between the output of the system and the output of the model:

$$\min_{\mathbf{p}} J_{mse}(\mathbf{u}(t), \mathbf{p}) = \frac{1}{t_{end}} \int_{t=0}^{t_{end}} \mathbf{e}(t)^T \cdot \mathbf{Q}(t) \cdot \mathbf{e}(t) dt \tag{3.4}$$

$$\mathbf{e}(t) = \tilde{\mathbf{y}}(\mathbf{u}(t)) - \mathbf{y}(\mathbf{u}(t), \mathbf{p}) \tag{3.5}$$

where $\tilde{\mathbf{y}}(\mathbf{u}(t))$ is the output vector of the process variables for a certain $\mathbf{u}[t : t_{end}]$ input profile, and $\mathbf{y}(\mathbf{u}(t))$ is the output of process model for the same $\mathbf{u}[t : t_{end}]$ input profile with $\mathbf{p}$ parameters. $\mathbf{Q}(t)$ is a user supplied square weighting matrix.

The classical optimal design criterion aims the minimization of a scalar function of $\mathbf{F}$ matrix. Several criterion exist:

- D-optimal experimental design maximizes the determinant of the Fisher matrix (Eq. 3.1), and thus maximizes the volume of the joint confidence region.

$$J_D = \max_{\mathbf{u}[t:t_{end}]} (det(\mathbf{F})) \tag{3.6}$$

- E-optimal experimental design is based on the ratio of the maximal and minimal eigenvalues of the Fisher matrix (Eq3.1). In ideal case this ratio is approximately one.

$$J_E = \min_{\mathbf{u}[t:t_{end}]} (\frac{\lambda_{max}}{\lambda_{min}}) \tag{3.7}$$

3.1.2 Calculation of sensitivities

Fisher information matrix is based on parameter sensitivities. In the following section the most common methods for the calculation of sensitivities, like direct differentiation method and finite difference method are described.

Sensitivities extracted from model equations

The analytical approach is the most accurate method to calculate the gradients. Consider the class of process models defined by Eq. 3.2 - 3.3, differentiate the state equation Eq. 3.2 respect to the model parameters, $\mathbf{p}$, and then integrate the resulted sensitivity equation on the considered time scale:

$$\dot{\mathbf{x}}_I(t) = \int_0^t \frac{\partial}{\partial \mathbf{p}} \frac{\partial \mathbf{x}(t)}{dt} dt = \int_0^t \frac{\partial}{\partial \mathbf{p}} f(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) dt \tag{3.8}$$

$$\frac{d\mathbf{y}(t)}{d\mathbf{p}} = \frac{d}{d\mathbf{p}} g(f(\mathbf{x}(t), \mathbf{p})) \cdot \frac{d}{d\mathbf{p}} f(\mathbf{x}(t)) \tag{3.9}$$

Eq. 3.8 can be solved simultaneously with the Eq. 3.2 state equation. The drawback of this methodology is its limited applicability due to difficulties in analytical differentiation of complex model equations.

The second approach integrates the Eq. 3.2 state equation first and then differentiates it respect to model parameters, $\mathbf{p}$. It is basically the first step to the numerical approximation of the sensitivities.

$$\dot{\mathbf{x}}_{II}(t) = \frac{\partial}{\partial \mathbf{p}} \int_{t-t_{sim}}^t \frac{\partial \mathbf{x}(t)}{dt} dt = \frac{\partial}{\partial \mathbf{p}} \int_{t-t_{sim}}^t f(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) dt \tag{3.10}$$

$$\frac{\partial \mathbf{y}(t)}{\partial \mathbf{p}} = g(\dot{\mathbf{x}}_{II}(t)) \tag{3.11}$$

Sensitivities calculated by finite difference of simulation results

Finite difference method is the most commonly used approach to calculate the sensitivities. This method is based on the finite difference approximation of the derivative from the solved differential equation.

$$\frac{\partial \mathbf{y}}{\partial \mathbf{p}_i} \approx \frac{\mathbf{y}((1+\Delta)\mathbf{p}_i) - \mathbf{y}(\mathbf{p}_i)}{\Delta \mathbf{p}_i} \quad i = 1, \ldots, n_p \tag{3.12}$$

where $\Delta \mathbf{p}_i$ is a small increment in the parameter value $\mathbf{p}_i$ and n_p is the number of the estimated parameters. Similarly to the previous case the length of time horizon (t_{sim}), in which the model equations are solved, highly influences the value of numerator of Eq. 3.12. As this approach is utilized in discrete form, t_{sim} means the number of samples which are included in the calculation of sensitivities.

When all the sensitivities are estimated at discrete time instants, calculation of Fisher information matrix is the following:

$$\mathbf{F} = \frac{1}{N} \sum_{i=1}^{N} \frac{\partial \mathbf{y}}{\partial \mathbf{p}}\Big|_{p=p^0}^{T}(i) \cdot \frac{\partial \mathbf{y}}{\partial \mathbf{p}}\Big|_{p=p^0}(i) \tag{3.13}$$

where N is the number of samples in the time horizon of the experiment (N=t_{sim}/t_{sample}).

This popular methodology has some drawbacks: the determination of the gradients by small perturbations of the parameters may give wrong results when these perturbations are too large - the approximation is no longer valid - or too small.

3.1.3 Time-series segmentation for supporting parameter estimation

The aim in this chapter is to describe the newly developed time-series segmentation method to support parameter estimation by extracting subsets of process data with high information content. In classical time-series segmentation univariate signal is analyzed. In wider interpretation the calculated sensitivities can be considered as multivariate time-series. It is very same as in dPCA based time-series segmentation framework (detect changes in the correlation structure of multivariate process data [60, 61]), where a set of dPCA models are calculated in each sample time. In Fisher matrix based approach we need to calculate sensitivities in each sample time. Dealing with parameter sensitivities we can get information about changes having effect in the parameter space of the model. The Fisher information matrix represents the correlation of the sensitivities. Instead of using either D- or E-optimality as the criterion for direct comparison of information content of two input sequences, we propose the use of Krzanowski's similarity coefficient between the eigenvectors of the Fisher information matrices obtained from the sequences. Thus, whereas the established criteria essentially compare the similarity of shapes or volumes of confidence regions for model parameters derived from the sequences, this new proposal also focuses on the similarity of their orientations.

The algorithm is based on the standard bottom-up scheme widely applied in off-line analysis of process data. In the following this segmentation algorithm will be presented.

Bottom-up time series segmentation

In the classical, univariate time-series segmentation the data set is mostly recorded by sensors, but in wider interpretation the calculated set of sensitivities could also be considered as a time-series. That is why the Fisher information matrix and an information criteria can also be calculated in each segments.

The basis for calculation of Fisher matrices is the sensitivity calculation of the model output respect to changes in model parameters. Thanks to discrete data acquisition, all the partial derivatives can be calculated in every sample time. In the considered time horizon, N (see Eq3.13), a set of partial derivatives shall be applied to calculate the Fisher matrix of an input signal sequence. Based on the Fisher information matrices D or E-criteria (Eq3.6 and 3.7 respectively) can be calculated to measure the information content of the considered input sequence.

Similarity of Fisher matrices

Information content of two different input sequences can be compared either based on the D or E-criteria or the direct comparison of Fisher information matrices. Fisher information matrix possesses superior information to criterion, since beside the quantity of information content it shows the direction of the examined information in the parameter space.

The similarity of Fisher matrices can be evaluated with using the Krzanowksi similarity measure ([27, 28]). It is developed to analyze similarities of principal component analysis (PCA) models (hyperplanes). Fisher matrix can be assumed like the covariance matrix of PCA. It concerns model parameters, rather than data covariances, so the end results relate to the parameter space rather than the data space. As eigenvalues and eigenvectors of the Fisher matrix can be calculated, it makes this tool fit for the purpose of extracting segments to find maximal information content.

Fisher matrix based time-series segmentation

The cost function of the time-series segmentation (Eq(2.1) in general cases, Eq(2.4) in this particular case which is the Krzanowksi similarity measure) can be minimized by dynamic programming (e.g. [52]). As it is unfortunately computationally intractable for many real data sets, usually one of the heuristic, most common approaches are followed, which are proposed in Section 3.1.3.

By calculating eigenvalues and eigenvectors of the Fisher information matrix, the direction of information in the information space is also taken into account to segmentation. The pseudocode of the algorithm is presented in Algorithm 3.1:

Algorithm 3.1 Fisher matrix based time-series segmentation for historical process data

Calculate the sensitivities in every sample times.

Define the initial segments (define a_i and b_i segment boundary indices).

Calculate the Fisher information matrix in the initial segments.

Calculate the cost of merging for each pair of segments:
$mergecost(i) = cost(a_i, b_{i+1})$.
The merge cost is based on sum of the differences (1- Krzanowski similarity (Eq2.4)) of the Fisher matrices of the segments.
The merge cost is calculated by:

$$cost(a_i, b_{i+1}) = \frac{b_i - a_i}{b_{i+1} - a_i}(1-Sim_{F(\{a_i,b_i\},\{a_i,b_{i+1}\})})+(\frac{b_{i+1} - a_{i+1}}{b_{i+1} - a_i})(1-Sim_{F(\{a_{i+1},b_{i+1}\},\{a_i,b_{i+1}\})}) \tag{3.14}$$

where $S_{F(\{a_i,b_i\},\{a_i,b_{i+1}\})}$ is the Krzanowski similarity measure of Fisher matrix calculated in the segment with the boarders of a_i, b_i and a_i, b_{i+1}, respectively. $S_{F(\{a_{i+1},b_{i+1}\},\{a_i,b_{i+1}\})}$ is calculated similarly using the segment with boarders of a_{i+1}, b_{i+1}.
while actual number of segments $<$ desired number of segments **do**
  Find the cheapest pair to merge:
  $i = argmin_i(mergecost(i))$
  Merge the two segments, update the a_i, b_i boundary indices, and recalculate the merge costs.
  $mergecost(i) = cost(a_i, b_{i+1})$
  $mergecost(i-1) = cost(a_{i-1}, b_i)$
end while

Merge cost formulated in Eq3.14 is basically the sum of differences of Fisher matrices in the "right-side" segment (with borders of a_i, b_i) to the recently merged segment (with borders of a_i, b_{i+1}) and the same for the "left-side" segment (with borders of a_{i+1}, b_{i+1}). There are weights formulated, for considering the length of the right-side and left-side segment and their contribution to the information quantity of newly merged segment. It is specially important when two segments should be merged with approximately equal length and a slightly different information content.

3.2 Case studies

The drawback of model based optimal experiment design is the relative high number of experimental runs which makes the parameter estimation costly and time consuming. Instead of optimization of the input trajectories the extraction of informative segments from historical process data can also support parameter estimation. This section will demonstrate the benefits of this concept by two application examples.

3.2.1 Segmentation of the input-output data of a first order process

First-order plus time delay models are widely applied in chemical process control, which is described by:

$$\tau \frac{dy(t)}{dt} + y(t) = Ku(t) \tag{3.15}$$

In this study, nominal parameters of the model are $K = 1$, $\tau = 10$ and the sample time is 0.1 sec.

Calculation of sensitivities

In this example, effects of differences in calculation ways of sensitivities are demonstrated. We investigated the dataset depicted on Figure 3.1.

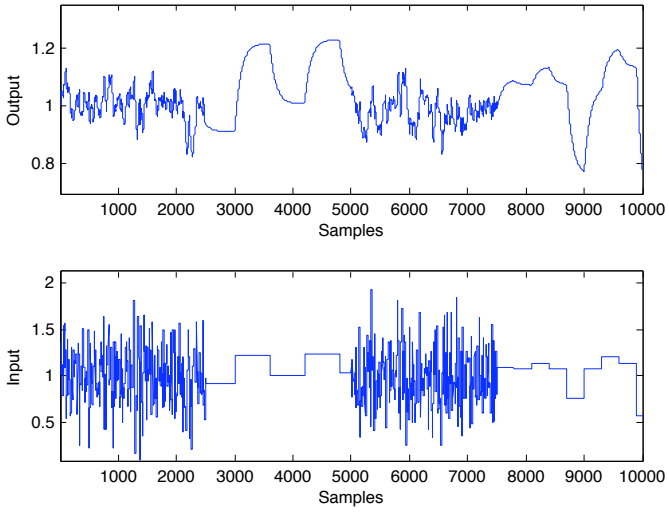

Figure 3.1: Input-output process data of the first order process

The analytical sensitivities can be calculated by differentiating Eq. 3.15 first and then solving the partial differential equation simultaneously with the state equations (I. approach):

$$\frac{d}{dt}\frac{dy}{dK}(t) = \frac{1}{\tau}u(t) - \frac{1}{\tau}\frac{dy(t)}{dt} \tag{3.16}$$

$$\frac{d}{dt}\frac{dy}{d\tau}(t) = -\frac{1}{\tau^2}u(t) + \frac{1}{\tau^2}y(t) - \frac{1}{\tau}\frac{dy(t)}{dt} \tag{3.17}$$

The second approach presented in Section 3.1.2 is based on the solution of the state equation (Eq. 3.15):

$$y(t) = \frac{K}{\tau}\frac{\int_0^t e^{\frac{t}{\tau}}u(t)dt + c}{e^{\frac{t}{\tau}}} \tag{3.18}$$

The direct differentiation of Eq3.18 provides the third way to calculate the sensitivities (II. approach):

$$\frac{\partial y(t)}{\partial K} = \frac{1}{\tau}\frac{\int_0^t e^{\frac{t}{\tau}}u(t)dt + c}{e^{\frac{t}{\tau}}} \tag{3.19}$$

$$\frac{\partial y(t)}{\partial \tau} = -\frac{K}{\tau^2}\frac{\int_0^t e^{\frac{t}{\tau}}u(t)dt + c}{e^{\frac{t}{\tau}}} + \frac{K}{\tau^3}\Big(te^{-\frac{t}{\tau}}\int_0^t e^{\frac{t}{\tau}}u(t)dt - e^{-\frac{t}{\tau}}\int_0^t te^{\frac{t}{\tau}}u(t)dt\Big) \tag{3.20}$$

The differences in the proposed methods are shown in Figure 3.2.

The numerical approximation of the sensitivities are not exactly identical to the analytical sensitivities (Figure 3.2). In this particular case of finite difference method the simulation time (t_{sim}) is 250. Since at the calculation of Fisher matrices normalized values of sensitivities are applied, the constant shift in the values of sensitivities does not affect the information content (see Eq. 3.1).

Time-series segmentation scenarios

We determined the optimal input signal of the process to provide a good background of comparison. Several studies deal with optimal design of identification experiments [62, 63, 64, 65, 33, 66]. In this study binary signal is chosen to determine the optimal input signal for model parameter identification. OED tools are highly suitable for compute the optimal value of periodic time of the input signal. Figure 3.3 shows the information content of input signal (measured by E

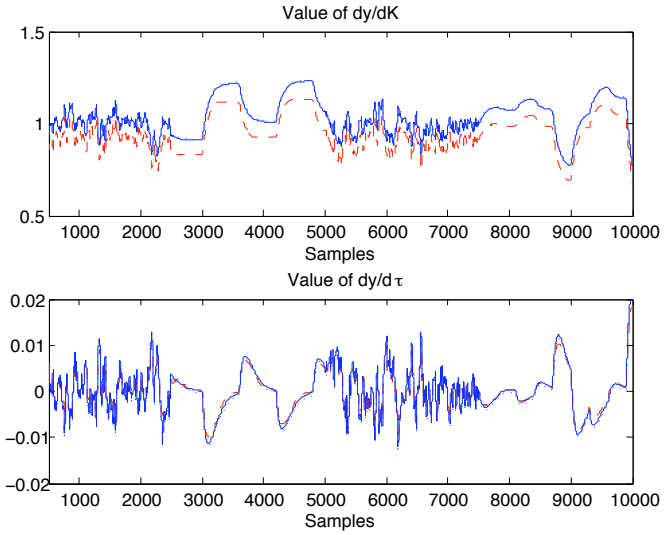

Figure 3.2: Sensitivities of the first order process. Comparison of thee calculation methods. full line - analytical sensitivity using I. approach, dotted line - analytical sensitivity using II. approach, dashed line - finite difference method

and D metrics) as function of different periodic times. As it is depicted, input signal with periodic time of four has the highest information content. This periodic time is almost the third of the time constant of the considered transfer function.

Based on the information yielded from Figure 3.3, we created a new dataset for segmentation scenario where sequences of "optimal" input signal are inserted. We calculated the set of sensitivity matrices and the proposed bottom-up time-series segmentation algorithm (Algorithm 3.1) is applied to find segments with different information content. In the Fisher information based time-series segmentation methodology the most important eigenvectors of the Fisher information matrices are compared directly using the Krzanowski similarity measure (Eq**??**). The eigenvectors with zero eigenvalues are excluded from the calculation. Initial segments are constituted of 2000 samples as minimal resolution. The number of desired segments was set to ten as stopping criteria of the segmentation algorithm.

As expected result, the dataset with low frequency constitutes an entire segment and segregated from the data set segments with high frequency. The segment with periodic time of 33% of the time constant is also automatically detected. This segment has the highest information content, depicted in Figure 3.4.

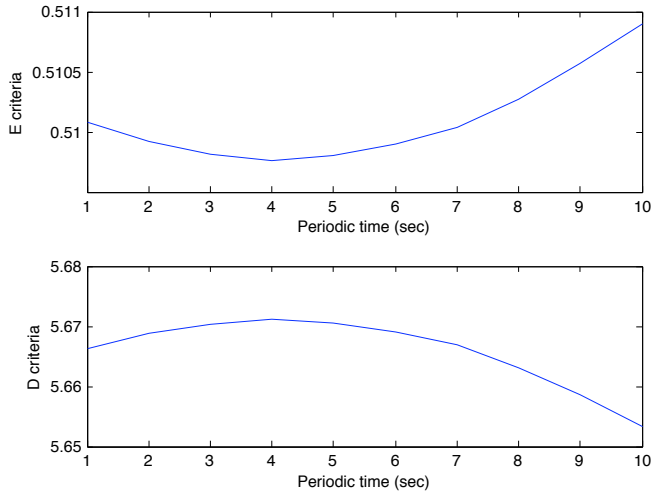

Figure 3.3: Information content of binary signals with different periodic times

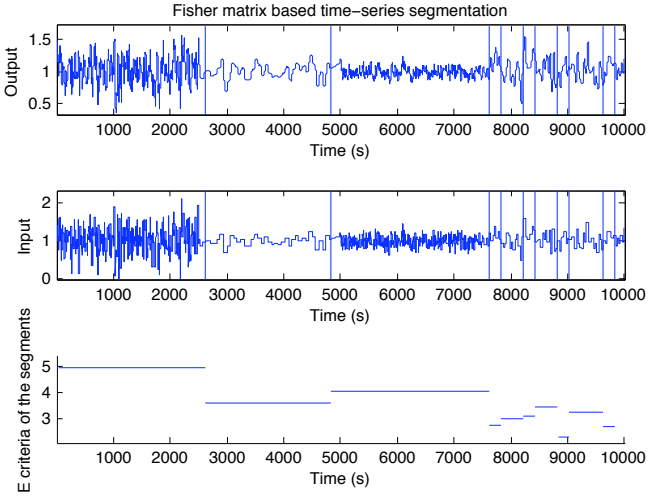

Figure 3.4: Result of Fisher matrix based time-series segmentation and the value of E criteria in the segments are also shown.

Identification scenarios based on the results of time-series scenario

To prove the benefit of Fisher matrix based time-series segmentation method, three different identification scenarios are considered:

- identification using optimal input signal sequence with the periodic time of 4.
- identification scenario based on the result of performed time-series segmentation using the segment with the highest E criteria value. It is for representing the segment with the lowest information content. In this example, it is the first segment.
- identification scenario based on the result of performed time-series segmentation using the segment with the lowest E criteria value. It is for representing the segment with the highest information content. In this example, it is the second segment.

The identification scenarios are performed using MATLAB and its fmincon function. The initial condition for K=1.5, time constant = 14. Parameters of the transfer function could be determined in all scenarios, only the number of function evaluations of optimization (minimizing Eq3.4) was different. Using input data sequence with the optimized information content, the lowest number of function evaluations was needed (43 iterations). Using the result of segmentation algorithm, the segment with high information content can provide almost the computational demand (44 iterations). As it can be expected the first segment with the less conditioned information matrix required the largest number of iterations (51 iterations).

3.2.2 Example with synthetic data of a polymerization process

Identification of highly nonlinear process models is more complex task than the previously presented illustrative example. Due to the complex nonlinear effect of parameters it is really necessary to support the parameter estimation procedure by information rich data regarding to the estimated parameters. Polymerization processes and their first principle models are highly suitable for representing the characteristics of nonlinear process models. The task is to automatically determine information rich segments that are applicable to the identify the parameters of the white box model described in the following subsection.

Description of the process

A continuously stirred tank reactor (CSTR)is considered in which a free radical polymerization reaction of methyl-metacrylate using azo-bis-isobutironitril (AIBN) as initiator and toulene as solvent. The number-average molecular weight (NAMW) is used for qualifying the product and process state. The polymerization process can be described by the following model equations, [67]:

$$\frac{dC_m}{dt} = -(k_p + k_{fm})C_m P_0 + \frac{F(C_{min} - C_m)}{V} \tag{3.21}$$

$$\frac{dC_I}{dt} = -k_I C_I + \frac{F_I C_{Iin} - FC_I)}{V} \tag{3.22}$$

$$\frac{dT}{dt} = \frac{(-\Delta H)k_p C_m}{\rho C_p} P_0 - \frac{UA}{\rho C_p V}(T - T_j) + \frac{F(T_{in} - T)}{V} \tag{3.23}$$

$$\frac{dD_0}{dt} = (0.5k_{tc} + k_{td})P_0^2 + k_{fm}C_m P_0 - \frac{FD_0}{V} \tag{3.24}$$

$$\frac{dD_1}{dt} = M_m(k_p + k_{fm})C_m P_0 - \frac{FD_1}{V} \tag{3.25}$$

$$\frac{dT_j}{dt} = \frac{F_{cw}(T_{w0} - T_j)}{V_0} + \frac{UA}{\rho_w C_{pw} V}(T - T_j) \tag{3.26}$$

where

$$P_0 = \sqrt{\frac{2f^* C_I k_I}{k_{td} + k_{tc}}} \tag{3.27}$$

$$k_r = A_r e^{-E_r/RT}, \qquad r = p, fm, I, td, tc \tag{3.28}$$

The mathematical model of the simple input simple output process consists of four states (C_m, C_I, D_0, D_1) and four nonlinear differential equations, where the manipulated input is the inlet initiator flowrate and the output is the NAMW defined by the ratio of D_1/D_0.

The mathematical model of the multiple-input multiple-output process consists of six states (C_m, C_I, T, D_0, D_1, T_j) and six nonlinear differential equations. By assuming an isotherm operation mode the process model could be reduced to four differential equations by neglecting $Eq(3.23)$ and $Eq(3.26)$, which still yields a highly nonlinear process but an easier way to investigate the proposed methodology.

Time-series segmentation scenarios and results

The process data that shall be segmented is depicted in Figure 3.5.

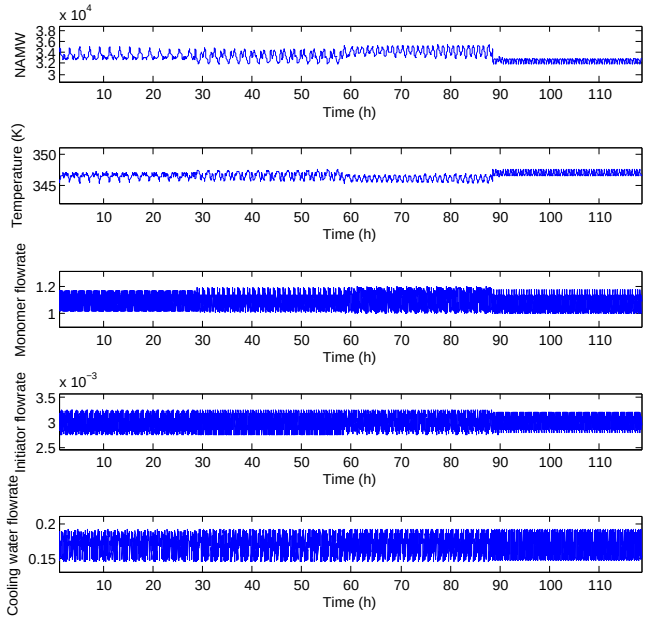

Figure 3.5: Process data used in the polymerization reactor example

Firstly all kinetic parameters of the model defined in $Eq(3.28)$ are considered unknown and involved in the segmentation (identification) procedure.

As the state equations are built up as complex combinations of model parameters, the model parameter estimation is quite difficult. Also due to this complexity the derivation of sensitivities from state equation is complicated, so the finite difference method is chosen to generate the sensitivities. The sample time in this case is $0.03h$. In sensitivity calculation, the simulation time for sensitivity calculation (t_{sim}) is chosen to be 100 time samples, which is longer than the dominant time constant of the process (which is almost $1h$). See the definition of simulation time in Eq. 3.12.

The first step of the segmentation procedure is the selection of the minimal resolution. In this particular case initial segments consist of 1000 samples. Using the bottom-up segmentation algorithm six segments were determined. The results of the time-series segmentation is summarized in Figure 3.6.

The first segment has the lowest information content and the fifth is the richest in this aspect. An identification procedure is performed using the data with the lowest and the highest informative segments to confirm this difference. In the identification

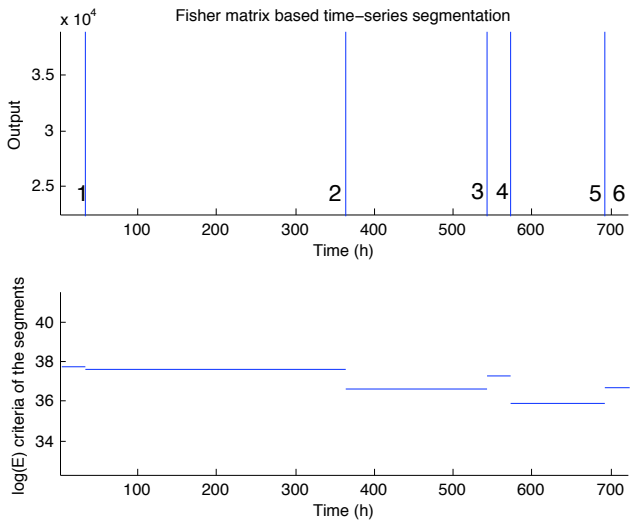

Figure 3.6: Result of segmentation for supporting to the identification of all kinetic parameters

process the following cost function is minimized:

$$\min_{A_r,E_r} \sum_{i=1}^{N} (100 \cdot (\tilde{y}_{i,T} - y_{i,T})^2 + (\tilde{y}_{i,NAMW} - y_{i,NAMW})^2) \tag{3.29}$$

where $\tilde{y_{i,T}}, y_{i,T}$ are output and calculated temperature values in i^{th} sample time, $y_{i,NAMW}\tilde{,} y_{i,NAMW}$ mean the same in terms of NAMW. The identification scenarios are performed using MATLAB and its fmincon function.

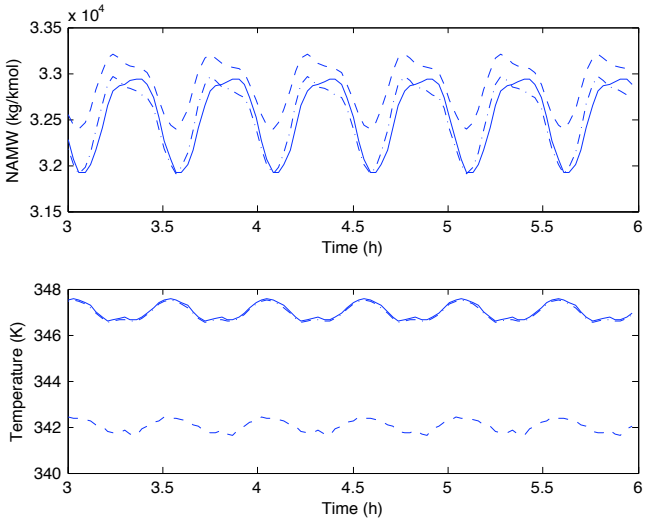

Figure 3.7: Results of identification scenarios (full line - original data, dashed line - the worst scenario, dashdot line - the best scenario)

Figure 3.7 shows informative results, the model identified from the segment with highest information content gives much better performance than model identified based on the worst segment. In this example all parameters were taken into account. Two further examples were designed to check the selectivity of the method respect to the parameter-set:

1. just the E_r parameters from Eq. 3.28 are considered as unknown and involved in the identification procedure.

2. just the A_r parameters from Eq. 3.28 are involved in the identification procedure. [1]

As first, just the determination of values of exponential parameters (E_r) is examined when preexponential (k_r) parameters are fixed on previously determined values. The segments with different information content are differentiated regarded to the exponential parameters. In Figure 3.10 the result of the segmentation is depicted with calculated the information content in each segments using the E criteria values. To be able to differentiate the segments they are marked with numbering.

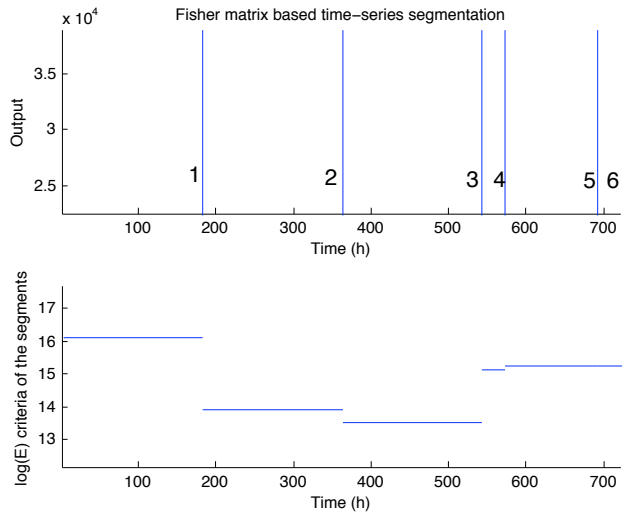

Figure 3.8: Result of segmentation for supporting the identification of exponential parameters

As it is depicted in Figure 3.8 the 3^{rd} segment has the highest and the 1^{st} has the lowest information content, respectively. Similarly to the previous scenario

[1] see results of segmentation scenarios in Appendix, Table A.3, Table A.4, Table A.5

where all the kinetic parameters are involved in the identification procedure a new identification process is performed to demonstrate the differences in information content. The results are summarized in Table A.4 and in Figure 3.9.

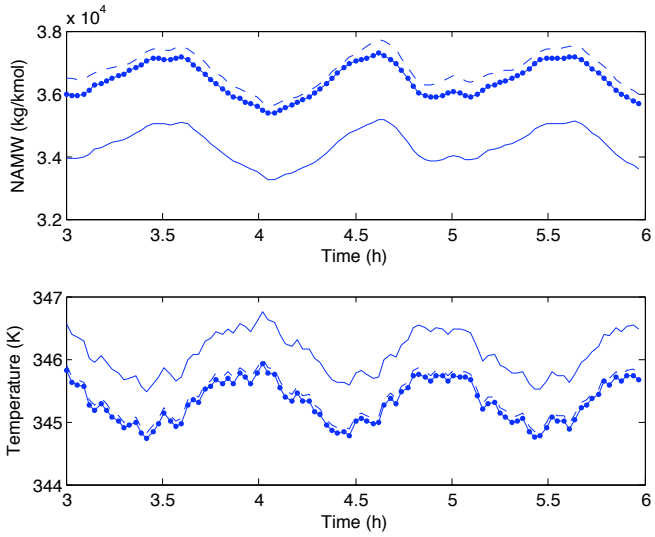

Figure 3.9: Result of identification of the exponential parameters (full line - original data, dotted line - best case, dashed line - worst case)

The value of the cost function is significantly reduced comparing the best case of this scenario and the previous case. This proves that some historical data segments have higher information content than the other ones.

The identification of the preexponential parameters (k_r) is examined to further improve prediction performance in the next scenario. In this case the exponential parameters are fixed in the value of the best case of the previous scenario. In Figure 3.10 the result of the segmentation is depicted with the information content in each segments using the E criteria.

As Figure 3.10 and Figure 3.8 show the result of the two scenarios are the same, but the information content of the segments are different as the identification point of view is changed. It shows that different segments of historical data are suitable for identification of different model parameters. Similarly to the scenarios above, an identification procedure is performed in this case too. The richest segment in information (related to the identification of the preexponential parameter) is 5^{th} and the poorest is the 2^{nd} as it is depicted in Figure 3.8. Results of this identification scenario is summarized in Table A.5 and Figure 3.11.

As Figure 3.11 shows the best case of the recent scenario approaches the original data quite well, since the difference is minimal (as it is shown in Table A.5).

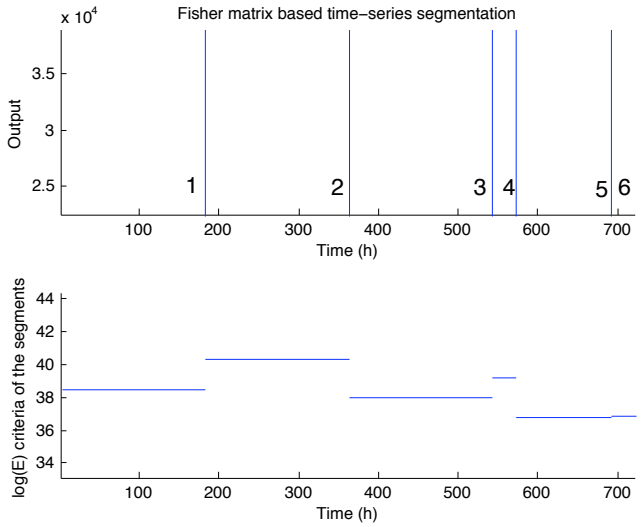

Figure 3.10: Result of segmentation for supporting the identification of preexponential parameters

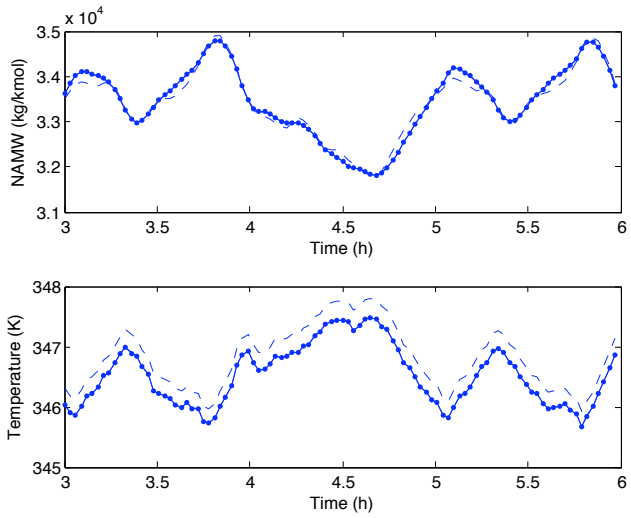

Figure 3.11: Identification result focusing to preexponential parameters (full line - original data, dotted line - best case, dashed line - worst case)

Table 3.1: Result of identification scenarios of determining all kinetic parameters

Scenario	Segment	$\frac{cost}{1000 samples}$	Type of parameter	r=p	r=fm	r=I	r=tc	r=td
Original	-	-	k_r	$1.77 \cdot 10^9$	$1.0067 \cdot 10^{15}$	$3.792 \cdot 10^{18}$	$3.8223 \cdot 10^{10}$	$3.1457 \cdot 10^{11}$
			E_r	$1.8283 \cdot 10^4$	$7.4478 \cdot 10^4$	$1.2877 \cdot 10^5$	$2.9442 \cdot 10^3$	$2.9442 \cdot 10^3$
Best	5^{th}	$7.71 \cdot 10^7$	k_r	$2.13 \cdot 10^9$	$1.75 \cdot 10^{15}$	$2.33 \cdot 10^{18}$	$2.53 \cdot 10^{10}$	$1.57 \cdot 10^{11}$
			E_r	$2.07 \cdot 10^4$	$8.16 \cdot 10^4$	$1.25 \cdot 10^5$	$3.95 \cdot 10^3$	$2.33 \cdot 10^3$
Worst	1^{st}	$3.38 \cdot 10^8$	k_r	$2.85 \cdot 10^9$	$9.09 \cdot 10^{14}$	$7.58 \cdot 10^{18}$	$4.42 \cdot 10^{10}$	$2.77 \cdot 10^{11}$
			E_r	$2.18 \cdot 10^4$	$7.97 \cdot 10^4$	$1.29 \cdot 10^5$	$2.11 \cdot 10^3$	$2.36 \cdot 10^3$

Table 3.2: Result of identification scenarios of determining exponential parameters

Scenario	Segment	$\frac{cost}{1000 samples}$	Type of parameter	r=p	r=fm	r=I	r=tc	r=td
Best	5^{th}	$4.0538 \cdot 10^3$	E_r	$1.8626 \cdot 10^4$	$7.4832 \cdot 10^4$	$1.2917 \cdot 10^5$	$5.7232 \cdot 10^3$	$3.4302 \cdot 10^3$
Worst	2^{nd}	$1.9573 \cdot 10^6$	E_r	$1.9978 \cdot 10^4$	$7.7165 \cdot 10^4$	$1.2775 \cdot 10^5$	$4.4439 \cdot 10^3$	$5.2659 \cdot 10^3$

Table 3.3: Result of identification scenarios of determining preexponential parameters

Scenario	Segment	$\frac{cost}{1000samples}$	Type of parameter	r=p	r=fm	r=I	r=tc	r=td
Best	5^{th}	165	k_r	$1.6281 \cdot 10^9$	$9.2605 \cdot 10^{14}$	$4.3787 \cdot 10^{18}$	$6.6971 \cdot 10^{10}$	$2.4946 \cdot 10^{11}$
Worst	2^{nd}	$1.6794 \cdot 10^7$	k_r	$2.1630 \cdot 10^9$	$1.6956 \cdot 10^{15}$	$2.3419 \cdot 10^{18}$	$2.6473 \cdot 10^{10}$	$1.5729 \cdot 10^{11}$

3.3 Summary of Fisher information based time-series segmentation methodology

In this chapter a novel time-series segmentation framework has been introduced to segregate segments from historical process data that are information rich in the parameter identification procedure. The methodology is based on the Fisher information matrix which possess the information content of a considered input signal. The information content of a data sequence can be measured utilizing D or E criteria.

The calculation of the Fisher information matrix is based on sensitivities of the model output respect to changes in model parameters. Some methods for calculation of sensitivities have been proposed in the chapter and their similarity has been investigated in details. The continuous calculation of parameter sensitivities makes the continuous calculation of Fisher matrices possible. This yields a time-series of Fisher matrices which provides the possibility to segment the original historical process data set based on their information content.

The Fisher information matrix possesses the quantity of the information and its the direction in the "information space", unlike to information criterion. To be able to evaluate the similarities of the Fisher matrices in the generated Fisher matrix time-series, Krzanowksi similarity measure is utilized, which is originally developed for comparing PCA subspaces. Integrating the Fisher information matrix and Krzanowksi similarity measure into the classical bottom-up time-series segmentation approach a novel tool is resulted, which can detect the changes in the direction of information in the "information space".

The applicability of Fisher information matrix based methodology is proposed throughout an example of simple input-simple output first order linear process and a more complex, multivariate polymerization example. In the latter example it has been proved that different segment are appropriate and information rich enough to estimate the whole set of parameters of the model and other segments can be segregated if just several parameters shall be estimated. In this example a detailed identification procedure can be followed based on the results of the time-series segmentation scenarios. Identification steps enhanced the assumption above, that some certain segments have more information content in the parameter determination point of view. In the final step of the whole identification scenario the difference of the original data and the the simulated data with the determined

parameters is minimal, which means that during the identification scenario the considered model parameters are well estimated in that certain operational point.

Chapter 4

Tuning method for model predictive controllers using experimental design techniques

Generally, one process is used for producing various products and satisfy various demands. So called off-specification products are produced during transitions between products. This product is generally worth less than the on-specification material (which fulfill all the commercial and quality requirements), therefore it is crucial to minimize its quantity. The on-specification product can be produced under varying circumstances and at varying operating points, which are more or less sound from an economical point of view, motivating the optimization of the production during production stages.

A large number of different grades are produced, and the transition times between the productions may be relatively long and costly in comparison with the total amount produced. The demand for reduction of the time and cost of grade transition inspires researchers to find innovative solutions [47, 48]. The optimization of complex operating processes generally begins with a detailed investigation of the process and its control system [31]. It is important to know, (i) how information stored in databases can support the optimization of product transition strategies, (ii) how hidden knowledge can be extracted from stored time-series, which can assure additional possibilities to reduce the amount of off-grade products. The optimization of product grade transition is a typical and highlighted task in process industry [49].

Advanced Process Control (APC) systems are designed to support the economic

operation both in process transients and in steady state operation. In most cases the operation of these control systems are based on a linear cost function, which usually contains the cost of the production and the price of raw materials and products. Obviously our goal is to maximize the quantity of on-specification materials and at the same time minimize the cost of the production by applying APCs. This is the top level of a multi-level optimization problem. As a second level of this problem it is inevitable to assure an appropriate and effective control strategy which is for realize the grade transitions and eliminate the effect of the disturbances as soon as possible. As Model Predictive Controllers (MPCs) are designed for handling these issues by defining an optimization problem, the application of MPCs is the lower level of the previously mentioned multilevel optimization problem.

Unfortunately, it is very difficult to find the right tuning parameters of the controllers in the whole operation range because of the nonlinearity of the process, and identified models (for MPCs) from input-output data are mostly linear.

Since these control systems are relatively expensive (limitedly accessible), the right parameters of the production (e.g. set-points, tuning parameters of controllers, valve positions) are determined experimentally using the intuition of engineers.

One of the common experimentation approaches is one-variable-at-a-time (OVAT) methodology, where one of the variables is varied while others are fixed. Such approach depends upon experience, guesswork and intuition. On the contrary, the statistical tools like design of experiments (DoE) permit the investigation of the process changing of factors-levels simultaneously using reduced number of experimental runs. Such approach plays an important role in designing and conducting experiments as well as analyzing and interpreting the data. These tools present a collection of mathematical and statistical methods that are applicable for modeling and optimization analysis in which a response or several responses of interest are influenced by various designed variables (factors) [50].

Modern optimal control and operation of a thermal plant and district heating network shall be a great project and the phenomena are highly similar to multiproduct chemical plants, especially if environmental aspects taken into consideration [68] and [69]. District heating networks (DHNs) could provide an efficient method for house and space heating by utilizing residual industrial waste heat. In such systems, heat is produced and/or thermally upgraded in a central plant and then distributed to the end users through a pipeline network. To reach environmental, operational and economical goals, proper and detailed description of the process is clearly needed like in [70] and [71]. Optimal operation means to

meet the consumers' and environmental requirements and at the same time fulfill the restrictions to make the operation of the plant safe.

Optimal control strategies meet these restrictions and at the same time minimize operational costs and environmental effects like described in Molyneaux's work [72]. Model predictive control (MPC) methods are highly applicable for these purposes since the formulation of the objective function might imply every aspects. The whole network has to be modeled, as MPCs require proper process model. The control strategies of these networks are rather difficult thanks to the non-linearity of the system and the strong interconnection between the controlled variables. That is why a non-linear model predictive controller (NMPC) could be applied to be able to fulfill the heat demand of the consumers.

The main objective of this section is to propose a tuning method for the applied NMPC to fulfill the control goal as soon as possible. The performance of the controller is characterized by an economic cost function based on pre-defined operation ranges. A methodology from the field of experiment design is applied to tune the model predictive controller to reach the best performance. The efficiency of the proposed methodology is proven throughout a case study of a simulated NMPC controlled DHN.

4.1 District heating networks as motivation example

District heating was promoted in Europe in the 1950s. Nowadays EU-CHP Directive could assure the legal framework for applying district heating for member states of the European Union. District heating network is implemented to utilize the heat generated by the combustion of city waste or industrial waste heat. Thanks to the efficiency and environmental friendly characteristics, the role of the district heating is still increasing [73]. The main advantages of district heating systems are the following:

1. Energy efficiency thanks to the simultaneous generation of heat and electricity in combined heat and power plants (CHPs).

2. Environment friendly by implementing renewable energy sources and utilizing industrial waste heat.

Several variations exist for district heating networks: the district heating network includes several consumers located in different areas like in [74], it can

contain an energy storage like in [75] or even lacks of thermal energy supply like in [76]. In some cases not just the local DHNs should be analyzed but the whole national DHN system, to investigate the sensitivity of the network to e.g. policy or even fuel price changes [77].

Model based control strategies (MPCs) are highly applicable for satisfying various control goals since the formulation of the objective function might imply every aspects. Model types of a district heating network in the literature can be a physical description of the heat and mass transfer in the network, like [78] and [79], and utilize node method like in [80]. There can be another approach, based on a statistical description of the transfer function from the supply point to the critical point considered. The forecast methodology proposed in [81] and [82] is to set an ensemble of ARMAX (auto-regressive moving average with exogenous input) models with different fixed time delays, and to switch between models depending on some estimated current time. In [83] the grey-box modeling approach combines physical knowledge with data-based, statistical modeling. Physical knowledge provides the main structure and statistical modeling provides details on structure and the actual coefficients/ estimates.

In this chapter the aim is to reduce the transition time in a non-linear model predictive controlled DHN by tuning the parameters of the non-linear MPC. The efficiency of the controller is measured by a cost function considering the limits of desired operation regime. To maximize this cost function the simplex method is applied, which is a well-known method in field of experiment design. This optimization method is able to handle mixed-integer optimization problems, which is needed because of the integer values of prediction and control horizon. Since there are periodic characteristics of heat demand, the proposed methodology can be easily inserted into an iterative learning control scheme ([84]).

The chapter is organized as follows: the topology of the district heating network will be described in Section 4.2. The applied MPC solution and the tuning method are introduced tn the second part of Section 4.2 and then control and optimization results will be examined in Section 4.3.

4.2 Modeling and control approach of a district heating network

4.2.1 The applied topology and modeling approach

The topology of the examined district heating network is presented in this section. The topology depicted in Figure 4.1 is chosen to represent the main characteristics of a district heating network. The network contains two heat production units, three consumers, two pumps and a valve. The production unit, called Producer 1, is the base load boiler, which may represent e.g. a waste incineration plant. The other production unit, called Producer 2, is the peak load boiler station, which has to satisfy the increased heat demand in the network, especially in case of Consumer 3. HX1 and HX2 heat exchangers are for transfer the produced heat from the primary circles to the secondary circle, which distributes the heat to the consumers directly.

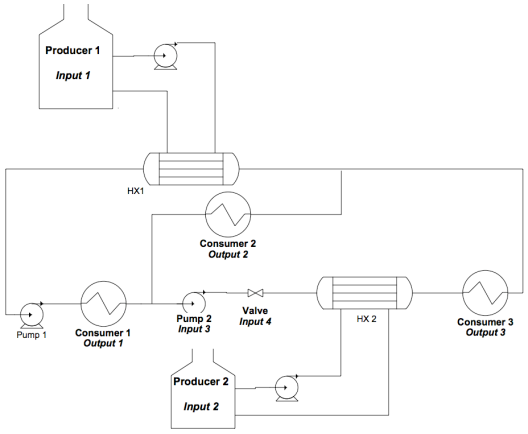

Figure 4.1: Topology of the district heating network

The model of this network is developed using the method of [78], which applies the physical description of the heat and mass transfer in the network. Structural approach is used to obtain a convenient global model: considering the complexity of the system, local models of the components of the network are established and then brought together.

Heat exchangers

In order to get the proper dynamic behavior of the heat exchangers an approach using a cell model with ordinary differential equations was chosen [10]. The heat exchanger was divided into perfectly and instantly mixed tanks, each featuring a

hot side and a cold side element (Figure 4.2). The logarithmic mean temperature difference of the heat exchanger is approximated more accurately as the number of cells increases. It is assumed that each cell is perfectly homogenous, and no back-mixing occurred. Also, the mixing is instantaneous. In our model five cells were used on the hot side and five cells on the cold side.

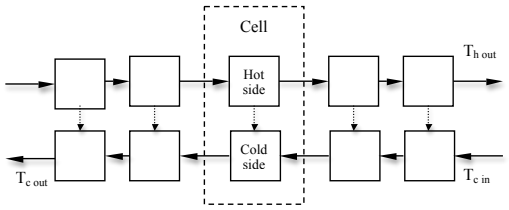

Figure 4.2: Cell model of the heat exchanger

The differential equations for the cells are shown in Eq(4.1) and Eq(4.2).

Hot side cell model:

$$\frac{dV_h \rho c_p T_h(i)}{dt} = V_h \rho c_p (T_h(i-1) - T_h(i)) \\ -UA(T_h(i) - T_c(i)) \tag{4.1}$$

where V_h is the volume of a cell on the hot side of the heat exchanger, ρ and c_p are the density and the heat capacity of the fluid, respectively, $T_h(i)$ and $T_c(i)$ are the temperature on the hot and cold side in the i^{th} cell of the heat exchanger, A is the area for heat transfer in a cell. To avoid the excessive complexity of the network the resistance of the wall is included to the heat transfer coefficient (U).

The cold side cell model is the following:

$$\frac{dV_c \rho c_p T_c(i)}{dt} = V_c \rho c_p (T_c(i+1) - T_c(i)) \\ +UA(T_h(i) - T_c(i)) \tag{4.2}$$

where V_c is the volume of a cell on the cold side of the heat exchanger, any further notations means the same as above.

Heat production units

The heat production units have been similarly modeled to the heat exchangers; the only difference is that only the cold side has been divided into cells.

Eq(4.3) represents the model of a cell (N is number of the cells, Q is the transferred heat):

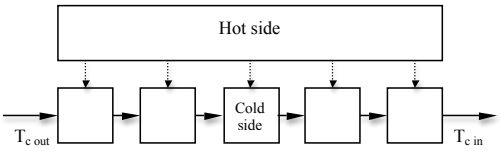

Figure 4.3: Cell model of the production unit

$$\frac{dV_c\rho c_p T_c(i)}{dt} = V_c\rho c_p(T_c(i-1) - T_c(i)) + \frac{Q}{N} \tag{4.3}$$

This simplification can be applied because since it is not important how the heat has been produced in the considered network just the quantity and distribution of the invested heat is significant. Detailed description of modeling approach a heat producer e.g. a CHP could be found in [85].

Pipelines

Pipeline network has significant effect on the operation of the considered DHN. This effect must be taken into account already in the piping network design ([86]). In order to model the pipelines of a DHN, two crucial effects have to be taken into consideration: the heat loss on the pipes could not be neglected. A more important factor, the dead time between both ends of the pipe must also be taken into consideration. The thermal energy propagation in pipes can then be modeled by a partial differential equation ([78]).

$$\frac{\partial T}{\partial t}(x,t) + \frac{m(t)}{\pi\rho R^2}\frac{\partial T}{\partial x}(x,t) + \frac{2U}{c_p\rho R^2}(T(x,t) - T_0) = 0 \tag{4.4}$$

where T is the temperature, m is the mass flow in the pipe, ρ is the density of the fluid in the pipe, R is the radius of the pipe, U is heat transfer coefficient on the wall and T_0 is the ambient temperature. This equation leads to the solution presented in Eq(4.5 ([78]):

$$T_{out}(t) = T_0 + (T_{in}(t - t_0(t)) - T_0) \cdot e^{-\frac{2U}{c_p\rho R^2}(t - t_0(t))} \tag{4.5}$$

As the thermal losses on pipes are assumed very low, the previous equation is

approximated by the following expression:

$$T_{out} \approx T_0 + (T_{in}(t - t_0(t)) - T_0) \cdot \left(1 - \frac{2U}{c_p \rho R^2}\right)$$
$$\approx T_{in}(t - t_0(t)) \cdot \left(1 - \frac{2U}{c_p \rho R^2}\right) \tag{4.6}$$

The computation of variable time delays is time consuming. That is why constant (and for instance nominal) time delays have been considered. This approach allows modeling thermal propagation as a simple non-linear dynamic system, which can be solved quickly.

The mechanical losses, pressure drop in pipes are modeled by:

$$\Delta p = \xi \frac{\rho v^2}{2} \frac{L}{2R} \tag{4.7}$$

where L is the length of the pipe, ξ is the mechanical loss coefficient and v is the velocity of the fluid in the pipe ($\frac{m}{sec}$) . Detailed modeling approach and description of the topology can be found in [87].

4.2.2 Multilayer optimization for DHNs

Tuning of an MPC to get better control performance can be considered as a multilayer optimization task. This problem is depicted in Figure 4.4.

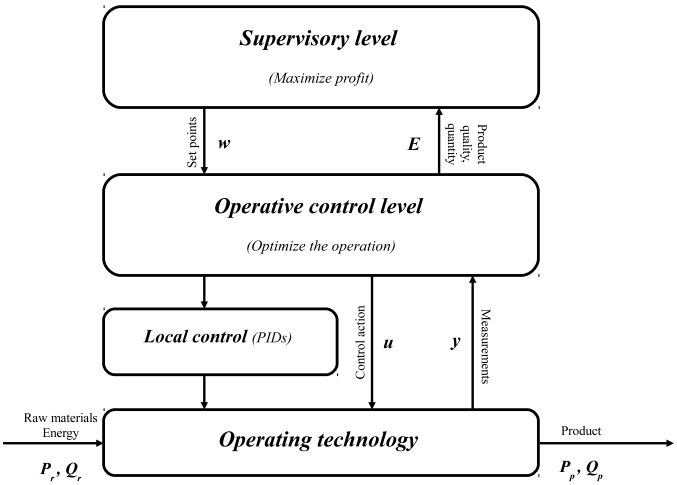

Figure 4.4: The layers of an economic optimization of an operating technology.

This framework is rather similar to the Advanced Process Control (APC) systems applied in process industry for online profit optimization [**?**].

The main goal is to minimize the transition time between two operation regimes from time to time. There is a need to formalize the problem first. The certain operation regime shall be specified by defining the upper and lower limits. The goal is to reduce the out-of-limits operation, which usually occurs during transitions. If the considered operation limits are not violated, the operation can be called appropriate. The linear cost function presented in Eq(4.8) is applied for measuring the appropriate operation time.

$$E = \sum_{i=1}^{N_o} P_i^{on} \cdot Q_i^{on} \tag{4.8}$$

where N_o is the number of consumers, P_i^{on} is the income (economic value in e.g. USD) when the consumed heat at the i^{th} consumer is inside the specification limits. Q_i^{on} means the consumed heat is between pre-determined limits (its value is 1 when inside 0 when outside the limits). This objective function shall be maximized by optimizing the tuning parameters of the applied model predictive controller. That is why there is a need to find a methodology which could handle mixed-integer optimization problems. The objective function (Eq(4.8)) represents the upper layer of the multilayer optimization problem. The MPC in the lower layer also formulates an optimization problem, hence the whole process optimization approach could be considered as a multilayer optimization problem.

4.2.3 Model predictive control of the DHN

Manipulated variables

The particular DHN described in Section 4.2.1 is considered in this section,. The possible manipulated variables are: the invested heat in Production unit 1 and 2, pump duty of P1 and P2 pumps and the valve opening. Since the P1 pump is chosen to compensate the pressure drop of the heat exchangers and pipelines, the P1 pump does not take part in satisfying the heat demand of consumers, so it was considered to be controlled by a local regulator.

The pressure drop in the direction of the Consumer 2 and in the direction of Consumer 3 must be the same. To reach this goal two manipulated variables can be used: the valve opening and the pump duty of the P2 pump. These manipulated variables are for determining the split ratio on the splitter and through this control

the flow in the two directions to be able to transfer enough heat to the consumers.

Analysis of applied models

Creating a mathematical model for control purposes is a challenging task in every MPC ([88]). In this case, since there were no real available operating plant, the process was replaced by the process model. This is based on the physical description of the DHN (called "A" model). "A" model is implemented in Simulink. In the examinations, a process model without time delays is going to be utilized for prediction ($t_0 = 0$ in Eq(4.5)-(4.6), but also based on the physical description of the DHN). This model is implemented in Matlab and called "B" model. In commercial MPCs, usually linear models have been applied for prediction (such as Dynamic Matrix Controllers, see [89]). It is necessary to update the model parameters regularly to keep the model valid in every operation range due to the nonlinearity of the controlled system.

The application of two different models has an important advantage: it is possible to simulate the situation when the model is not able to describe the operating process perfectly. A non-linear model, based on the physical description of the system is created to reduce the necessity of updating the model parameters and extend the validity of the model in the whole operation range. The prediction ability of the model is based on the "measurements" of the controlled system, which are applied for parameter estimation purposes. The difference of the "operating network" and the process model for prediction is caused by assuming a different time delays as described previously.

To demonstrate the differences between the "operating network" and the process model used in the MPC an examination has been carried out. The results of the comparison shown in Figure 4.5 with respect to the same input signals.

Objective function and constraints of the model predictive controller

The first task is to define the possible manipulated variables, when creating the model predictive control system of a district heating network. These variables can be either continuous ([90]) or integer variables (e.g. boiler status ([91])). In case of optimization this leads to a mixed integer optimization problem. Solving an optimization problem like this is rather difficult, time consuming and computationally demanding. In this example a simple non-linear sequential quadratic programming (SQP) method with soft constraints will be applied to avoid

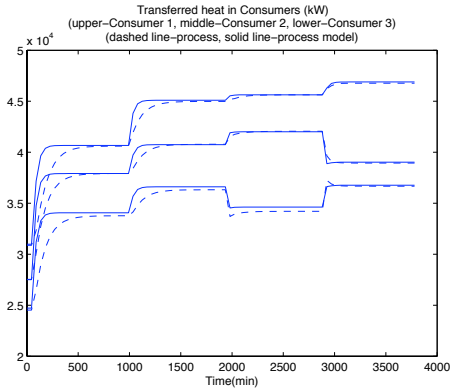

Figure 4.5: The outputs of the "operating process" and process model to the same input signal

the difficulty of mixed integer non-linear programming (for more details see [87]). The solution to avoid the problem of mixed-integer optimization is to augment the conventional objective function of MPC with the absolute values of manipulated variables. To differentiate the importance of the manipulated variables different weights shall be applied for them in the extended objective function (e.g. utilizing heat invested from the base load boiler rather than applying the peak load boiler). The objective function of the utilized MPC is formalized in Eq(4.9).

$$\min_{\mathbf{u}(k+j)} \beta \sum_{j=1}^{p} (\mathbf{w}(k+j) - \mathbf{y}(k+j))^2 + \alpha \sum_{j=1}^{p} \mathbf{u}(k+j)^2 + \gamma \sum_{j=1}^{c} \Delta \mathbf{u}^2(k+j-1) \tag{4.9}$$

where $\mathbf{w}$ is the setpoint signal, $\mathbf{y}$ is the controlled variable, $\mathbf{u}$ and $\Delta\mathbf{u}$ is the absolute value and the change of the manipulated variable, $p, c, \alpha, \beta, \gamma$ are the tuning parameters of the MPC. The aim of the controller to fulfill the heat demand of consumers. $\mathbf{y}$ means the transferred heat in the consumers, calculated based on the difference of the outlet and inlet temperature of consumers on the cold side, Eq(4.2). The control goal is reached by varying the implemented heat in the production units. The transferred heat in the production units are symbolized by $\mathbf{u}$ the same as denoted with Q in Eq(4.3). The performance of the controller highly depends on its' tuning parameters and the forecast of the heat demands. So the determination of values of tuning parameters is crucial project in reduction of transition time.

In the case study, α is a vector with four elements: the weight for Producer 1 is 0,

since it is not necessary to punish the control actions of Producer 1. On the contrary the weight of the control action for Producer 2 is non-zero, since it is important to punish its' control action, utilizing the heat sources in Producer 1 instead. The situation is the same in case of the valve and the P2 pump since the control action of the valve is preferred to the control action of P2 pump. γ is a constant for punishing the change of valve position.

In the created MPC framework SQP optimization method has been utilized to minimize the objective function presented in Eq(4.9). The optimization in the MPC has to be realized taking into account the constraints of the process. These constraints express that the actuators have a limited field of action as well as determined slew rate, as in the case of valves. The input constraints in this study are formalized as in Eq(4.10).

$$\mathbf{u}(k+j-1) - \Delta\mathbf{u}_{max} \leq \mathbf{u} \leq \mathbf{u}(k+j-1) + \Delta\mathbf{u}_{max}$$
$$j = 1 \dots c \tag{4.10}$$

where c is the length of the control horizon.

Application of Internal Model Control scheme

There is an obvious model mismatch, shown in Figure 4.5. There are differences between process outputs and model outputs both in transients and in steady state. This mismatch motivated us to apply the Internal Model Control (IMC) scheme ([92]), depicted in Figure 4.6.

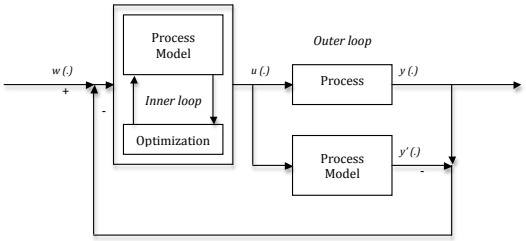

Figure 4.6: The scheme of the implemented non-linear model predictive controller

The IMC scheme is used in a modified form as follows. The IMC structure modifies the set point signals during transitions significantly which leads to huge overshoot during setpoint changes. Hence, it is not advantageous to apply this scheme during the transitions. At the same time it is very useful to apply the

IMC scheme to eliminate the steady state offset. So a trigger is implemented in the optimization box to switch on and switch off the IMC scheme. The trigger is formulated with the following expression:

$$\frac{\sum_{k=1}^{E} ME_k(i) - ME_k(i-1)}{E} \leq K \tag{4.11}$$

where ME is the modeling error vector in i^{th} and $(i-1)^{th}$ sample time, E is length of the modeling error vector, K is a constant. When the change of the model error is smaller than a previously determined constant, the controlled variable is relatively close to the set point. If this condition is fulfilled the IMC scheme will be expected to switch on and eliminate the steady state offset.

4.2.4 Methodology for tuning parameter optimization

There are no common practices to determine the parameters especially in case of punishment factors in the field of MPC tuning . A methodology has been introduced In [93] to solve this problem based on the first order and dead time models of the controlled object, but this method is valid only for linear MPCs.

Finding optimal tuning parameters might have numerous approaches: (i) based on pure engineering experience, (ii) using goal oriented optimization methods which can imply engineering experience, (iii) using pure simulation based optimization approach. Pure engineering experience is obviously not enough to find optimal tuning parameters, however essential to yield the parameter combination which assure safe operation. Approaches, which imply engineering experience, is the most transparent way of controller tuning, since control engineers can pair effect of changes in tuning parameters and their result. Pattern search methodologies can satisfy this need with determining the quantity and direction of changes in tuning parameters. Beside this, this tuning approach does not require detailed process model to find optimal parameter combination as simulation based optimization does. As MPC tuning is a mixed-integer optimization problem, it needs different tricks to find optimal tuning parameter combination: e.g. optimization in two rounds: determine optimal value of the integers first and then optimizing by the rest of parameters or using different relaxation methods. Both of the approaches leads to loose of transparency.

In our tuning approach the well-known simplex methodology was employed to maximize the objective function of Eq(4.8) with varying the tuning parameters of model predictive controller. This methodology is widely applied in field of

experimental design ([94]). The simplex method consist of the following steps:

1. In case of M pieces of variables, $M + 1$ pieces of experiments are necessary to be carried out to create the initial simplex. In this paper $M = 6$ since the following tuning parameters shall be adjusted: α (except the weight of control action of Producer 1), γ, p and c.

2. Evaluation of objective function at the peaks of the simplex. With reflecting the peak with the lowest value (since maximizing Eq (4.10)) to the opposite hyperplane defined by the residuary peaks, the parameters of the new experiment is found.

3. The obtained parameters is used instead of the reflected peak.

4. Carrying out the experiment with the new parameters.

5. Continue the reflection (Step 2 and Step 3) and determine the value of extension or contraction coefficient. Stop if the value of the objective function reach the desired value.

The equation for the procedure of reflection can be written:

$$\mathbf{x}_m^{new} = \frac{1+\lambda}{M} \cdot \sum_{i=1}^{M+1} \mathbf{x}_i - (\lambda + \frac{1+\lambda}{M}) \cdot \mathbf{x}_m \tag{4.12}$$

Where M is the number of optimized variables (the length of $\mathbf{x}$), $\mathbf{x}_i$ stands for the coordinates of the simplex before reflection, $\mathbf{x}_m$ is the parameters of the simplex with the worst value of objective function and λ is the extension or contraction coefficient.

The convergence of simplex methodology (or Nelder-Mead approach) has been examined in details by Lagarias et al. in [95]. They found this method might have failed to get the optimal solution even in simple, low dimensional optimization cases. The other remark in their investigation is a huge advantage of this pattern search method: it can decrease the value os the cost function very fast. Practically in controller tuning, there is no need to find the optimal solution, but finding a much better tuning parameter combination is much more important, which makes simplex methodology suitable for this purpose.

As pattern search algorithms do not need the gradient of the cost function respect to optimization parameters, most of them are suitable for this controller tuning approach. Kiraly et al. ([96]) used Mesh Adaptive Direct Search (MADS)

methodology ([97]) in finding optimal operation points for different controllers respect to pre-determined constraints (standard PI controller and an MPC). Full factorial plans of experimental design can also be applied for this reach the tuning goal. What is common in these methodologies: pattern search algorithms enable us to tune controllers with setting optimization parameters manually after evaluation of the cost function. Experience of control engineers can also be implied into the tuning process this way.

4.3 Results and discussion

In the case study the main goal is to maximize the cost function (formulated in Eq (4.8)) by tuning the MPC. We have to define low and high limits to highlight the range, where we consider the operation is fulfilling the requirements. In this particular case 1% of the setpoint is considered as acceptable distance from the desired operation value. The tuning parameters are values of the prediction and control horizons and values of α and γ in Eq (4.9). These parameters represents the search space where the simplex methodology is applied.

In the optimization scenario seven simulations have been executed to initialize the simplex, and five more to improve the control performance. Four more experiments were evaluated to prove that the optimum solution was reached. During these experiments the simplex seemed to rotate around, which indicates that the maximal value of objective function (Eq (4.8)) was reached.

The performance of the controller with initial and tuned parameters have been compared graphically first, in Figure 4.7 and Figure 4.8.

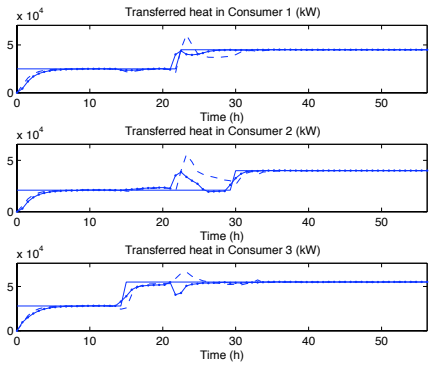

Figure 4.7: Comparison of the transitions in outputs with initial parameters (dashed line) and with the experimentally determined parameters (dotted line)

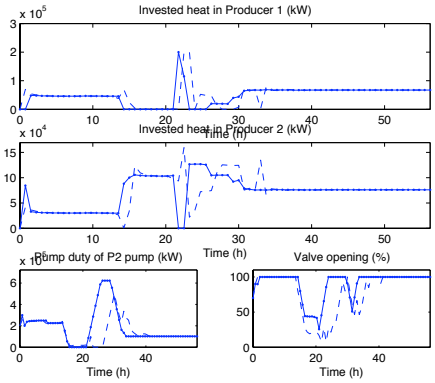

Figure 4.8: Comparison of the transitions in inputs with initial parameters (dashed line) and with the experimentally determined parameters (dotted line)

The time demand of the transition is shortened. Although a badly tuned MPC is also capable of performing the transition, the optimization of tuning parameters is necessary. The parameters of the MPC determine how effective the de-coupling of the process variables is, which is presented throughout the example of Consumer 2. In this case the regulator could not eliminate the effect if there is a transition at Consumer 1. In the optimized case the heating network fulfills the requirement of Consumer 1, 2 and 3 more than 63 %, 52% and 57% of the examined time horizon (it is the on-specification time in the considered period.) In case of initial guess, these ratios were 50%, 45% and 52%, respectively. It has another advantage as well, because the MPC tuned with the proposed method eliminates the effect of changes in the heat demand of other consumers, quite effectively.

As a second comparison, traditional Integral of Square Error (ISE) and Integral of Absolute Error (IAE) performance metrics are computed. These metrics aggregate difference between setpoint and controlled value over a considered time period. Hence these metrics can be considered stricter than Eq (4.8), since every perturbation from the setpoint has been taken counted. Values of the metrics and comparison of the optimized and initial tuning have been summarized in Table (4.1).

Table 4.1: ISE and IAE metrics for the considered control scenario

Metric	Tuning method	Consumer 1	Consumer 2	Consumer 3
ISE ($*10^9$)	Initial guess	1.75	4.61	2.82
ISE ($*10^9$)	Optimized	1.65	1.90	2.27
ISE (%)	$\frac{Optimal}{Initial}$	95	41	80
IAE ($*10^5$)	Initial guess	1.66	2.75	1.98
IAE ($*10^5$)	Optimized	1.28	1.70	1.83
IAE (%)	$\frac{Optimal}{Initial}$	77	62	92

By further investigation of Figure 4.8, less movement in manipulated variables can be noticed. It indicates smoother transient between operation points. To measure the total movement of each manipulated variable, the following expression has been defined:

$$IDU_n = \sum_{k=1}^{N} abs(\Delta u_k) \tag{4.13}$$

where IDU_n means the integral of absolute value of the Δu denoted by n over the considered time horizon ($k = 1, ..., N$). The following result can be summarized in Table (4.2), based on this metric:

Table 4.2: Total movement of manipulated variables (IDU) in the considered control scenarios

Tuning method	Producer 1	Producer 2	Pump duty	Valve Opening
Initial guess	9.5*10^5	7.5*10^5	1.74*10^6	500
Optimized	5.7*10^5	5.1*10^5	1.65*10^5	300

Based on the results above, the following statements can be made:

- a cost function (like Eq (4.8)) or a performance metric (like ISE or IAE) shall be defined in the first step to measure MPC controller performance.
- simplex methodology is highly applicable to improve the value of the defined cost function by varying MPC parameters simultaneously. It needs low number of iteration for significant improvement.
- all the cost function and performance metrics showed remarkable improvement in control performance by tuning the controller this method.
- quicker transition between operating points, quicker disturbance elimination can be noticed.
- less movement in the manipulated variables which indicates smoother/ more robust control actions.

4.4 Conclusion

A multilevel optimization approach of a district heating network has been presented in this chapter. The main goal was to fulfill the heat demand of consumers as soon as possible in a non-linear model predictive controlled DHN. To reach the goal

of shortening the length of transients the first task is to create a cost function for measuring the efficiency of control. This cost function is based on the income if the consumers' heat demands are fulfilled. The next step is installation of the non-linear model predictive controller (NMPC). The model applied for prediction is based on the physical description of heat and mass transfer. The internal model control (IMC) scheme has been utilized to take the possible model error into consideration. The optimal tuning parameter combination of the NMPC provides the shortest transient time and the maximal income. The simplex method can be a good choice to find these parameters as this method involves reduced number of experimental runs to localize the optimal value of tuning parameters. The efficiency of the proposed methodology has been shown by a case study where the transition time is decreased by 10% .

Chapter 5

Summary

Methods of technology development become wider and wider, thanks to the information potential in enormous amount of archived process data. The mathematical modeling and optimization approaches are used more and more commonly beside the intuition and experience based tools. Beside the statistical tools - e.g. SPC (Statistical Process Control), PCA (Principal Component Analysis) based fault detection - more complex tools, which are based on the explicit utilization of physical-chemical laws can assure the wide-spread assistance to solve engineering issues. There is one common characteristic in all approaches: successful application is based on the extraction of useful information from the data storage and at the same time considering the relationship between input-output process data, which means creating a proper process model.

The first step of every model based process development after the detailed investigation of the considered process is the determination of followed modeling method and creating the appropriate model structure. The model structure depends on the purpose: e.g. the APC (Advanced Process Control) tools use linear, black-box models, the OTS (Operator Training System) tools use non-linear first principle models to describe the behavior of the process. The next step of the modeling process is to estimate the model parameters which are based on the proper selection of the input-output data slices. It is rather important that effects of misbehavior and process faults have to be removed and periods of linear or non-linear operation ranges have to be segregated. The selected data sets are used in every step of model parameter estimation from identification to validation as well as they might have huge impact on the results of the economic studies during the last steps in a model based process development.

The aim of the thesis is to introduce some novel and innovative tools to

segregate the process data in order to: (i) detect changes in linearity in input-output correlation, (ii) help to select informative segments in model parameters estimation process. Beside these tools an aspect of economic based process development method is also investigated.

Locally linear correlation structure of input-output data can be changed by faults, process misbehaves or switching operation point. The commercial industrial fault detection tools are mainly based on PCA. As a first step during application of these tools, a fault-free operation period is selected to create the PCA model. Using this PCA model, the difference of the recently collected process data and the predicted process value by the PCA model is computed and determined if a possible fault occurred. In case of using the proposed dPCA based time-series segmentation methodology, we can extend our field of interest, during data-processing. There might be double types of goals: (i) fault detection and (ii) detecting the changes in the linear relationship of input-output process data. They sound similar but there are important differences. By detecting changes in the correlation structure of input-output data, we are interested in finding time-periods in which different linear relationship of input-output process data is valid. This information can be highly appreciated and very well usable in the field of Advanced Process Control applications. Goal oriented applications are developed for segmentation of historical and streaming process data which widen the possible field of utilization.

After segregating the data of fault-free operation ranges, one of the first tasks is the model building and model parameter estimation. Proper selection of operation periods is inevitable to find time-frames with high information content to support parameter estimation. To handle this problem the Fisher information matrix can be a very powerful tool from Optimal Experiment Design (OED) toolbox. This matrix contains the sensitivities of model output, which is basically the partial derivatives of model output respect to the model parameters while considering a given input sequence. Based on the Fisher information matrix, a novel, innovative time-series segmentation method has been proposed which helps to segregate the operation periods with high information content in the model parameter estimation process with a pre-defined model structure. T the same time the information content can be measured by E and D criterias from OED toolbox.

The other type of experimental design is the classical design of experiment methodology (DoE) which can be effective applied in data-driven, economic oriented process development. In this field, the Advanced Process Control (APC) applications became wide-spread, which have the basis of Model Predictive

Controllers (MPCs). The performance of these controllers highly depends on the applied tuning parameters beside the prediction ability of the applied process models. These tuning parameters have huge effect on the economic performance, which can be measured by a goal oriented objective function. The economic optimization by varying the tuning parameters is a mixed-integer optimization problem. A new framework has been created to systematically get closer to the economic optimum with considering the bottlenecks and operation limits of the process. If the detailed process model of the considered process exists, it is possible to determine the optimal tuning parameters in the design phase of the controller. If not, the framework can be integrated into the iterative learning control scheme, which provides the possibility to get closer to the economic optimum step-by-step from one product cycle to an other.

5.1 New Scientific Results

1. **Off-line and on-line time series segmentation algorithms were developed with utilizing Dynamic Principal Component Analysis and recursive covariance matrix computation to segregate homogeneous operating ranges and detect faults, process misbehaves or operating point changes.** (Related publications: 4, 7, 12)

   A potential possibility to improve operating technologies is to detect homogeneous operation regimes and detect the occurrence of faults and misbehaves, which might break the homogeneity. The only thing that is given in this case is the process data and the assumption in which the linear relationship of input-output data is supposed. This approach is one of the key approaches in the application of model predictive controllers since these are based on linear models, so validity of the models can be determined.

   In analysis of correlation of process variables in multivariate data sets principal component analysis (PCA) is wide-spread applied. Since time-dependency of process data is not taken into account in traditional PCA, dynamic PCA is applied to handle this problem, where the data matrix - constituted from input-output process data - is augmented with the values collected in the previous sample times. As PCA is a statistical methodology, high quantity of data is necessary to be able to compute covariance matrices,

which is an obstacle in the accurate detection of the occurrence of faults. A new covariance matrix has to be computed in every sample time to improve the resolution. The recursive computation method of covariance matrices is applied where actual process samples and the covariance matrix computed in the previous sample times are used to calculate the recent covariance matrix. Hence, we get a time-series of covariance matrices as a result.

A key element of recursive computation is the forgetting factor, which is for weighting the recently collected process data against the previously computed covariance matrix. The effective, Fortescue et al. defined, variable forgetting factor is applied to assure the quick adaptation of the covariance matrices to recent operation range. The similarity of these matrices is defined by the Krzanowski measure in the time-series of covariance matrices, which is the cosine of the angle of two dPCA models, practically.

Off-line and on-line multivariate time-series methodologies are developed to detect the accurate time of the occurrence of faults and misbehavior in historical and streaming process data by integrating these tools into the classical bottom-up and sliding window segmentation techniques. The developed framework is tested and examined throughout the benchmark Tennessee Eastman process.

2. **Utilizing the tools of optimal experiment design - Fisher information matrix, D and E criterion - a novel time-series segmentation methodology has been developed in which the historical process data can be segmented to highlight time-frames with high information content regarding to parameter estimation process of a mathematical process model with pre-defined model structure.** (Related publications: 1, 5, 10, 11)

   As mathematical models of chemical processes become more and more wide-spread, there is a huge demand to predict the process behavior more and more effectively. The keystone of these solutions is to determine the appropriate model parameters in the considered operating range. In this model development step, we focus on the proper selection of input-output data slices. Two options are available to reach data slices with high information content: (i) to design and carry out proper experiments which are

time-consuming and cost demanding or (ii) the other way is to segregate these data sets from historical process data.

There are tools for determining the information content of a particular input-output data set. These tools are based on the Fisher information matrix which is constituted from partial derivatives of model outputs respect to model parameters (sensitivity equations) using a considered input data set. D and E criterion can be calculated to measure the information content of the considered input data set based on the Fisher information matrix.

The Fisher information matrix implies the information content of a given data set but at the meantime it can prove additional information about direction of the information content in the parameter space of considered set of model parameters. This helps to segregate the process data segments which have the same aggregate information content (calculated with D or E criterion), but the model parameters in the parameter set have different contribution to aggregate information content. This information is stored in the eigenvectors of Fisher information matrix similarly to the eigenvectors of the covariance matrix of PCA. With utilizing this feature, a novel time-series segmentation methodology has been developed in which the similarity of the Fisher matrices is determined with using the Krzanowski similarity measure. These tools have been integrated to the classical bottom-up time-series segmentation methodology to detect the changes the direction of information content in the information space set by model parameters.

3. **A methodology is developed based on the classical experiment design techniques which is effectively applicable for improving and optimizing the operation of the already installed or design-phase model predictive controllers.** (Related publications: 2, 3, 6, 8, 9, 13)

   In parallel with the development of the modern control systems there is a need to calculate and maximize the economic benefit by implementing the most recent control techniques. In the latest advanced control technologies, model predictive controller are wide-spread applied. Setting the tuning parameters of these controllers properly requires a highly experienced control engineer to achieve the highest economic performance. These tuning parameters are even more important in case of changing operating point to minimize the possible off-grade product.

The developed methodology is based on an economic objective function which has aim of either cost minimizing or benefit maximizing. This is also for measuring the control performance beside the applied controller tuning parameters. It is shown that by utilizing the simplex methodology of experiment design, the tuning parameter values of model predictive controllers can be optimized in spite of the mixed-integer optimization problem caused by the time horizons and suppression factors and at the meantime the operation limits can be effectively considered.

The developed methodology can be applied in various stages of controller design and development:

(a) 1. If the mathematical model of controlled process exists then is possible to calculate the economic performance in an operating point or grade change with a defined tuning parameter set using the control system - controlled object simulator. Integrating this simulator and economic cost function with the simplex methodology, it is possible to determine the tuning parameter set which provides the highest economic benefit in a considered scenario.

(b) 2. If the mathematical model of the operating process is not available then inserting the tools of classical experiment design and the economic cost function into the iterative learning control scheme, tuning parameters of the controllers can be set manually and the economic performance can be improved from one production cycle to another.

5.2 Future Work

The proposed algorithms and results have the potential to follow further, interesting research field.

- To support recently applied engineering techniques, integration of dPCA and Fisher based time-series segmentation method can be a huge advantage, e.g. in design phase of APC projects. It means the collected data from DCS can be first segmented by linearity point of view and then the resulted segments can be segregated using the Fisher information matrix based algorithm to

select the most informative periods. A user friendly implementation of this integration can be shorten the time-consumption of the step test phase in which the operating process is disturbed to be able to collect data for model parameter estimation.

- An interesting research field can be the utilization of Independent Component Analysis (ICA) and its dynamic version in fault detection and time-series segmentation. ICA can be used to describe the relationship between input-output process data instead of using PCA.

- A huge disadvantage of PCA based time-series segmentation is to lack of the ability to determine the reason of the occurred fault, just determining the time of occurrence. Similarly to the Artificial Immune based fault detection methods, it might be useful to investigate the possibility to extend the PCA based time-series segmentation to be able to detect the reason of the change in linearity e.g. using supervised learning techniques.

- Beside the time-series segmentation methods, the further investigation of simplex based optimization method can be also useful, since the possible application ways of the historical process data e.g. in the initialization phase, have not been examined in details.

Chapter 6

Összefoglaló

6.1. Bevezetés

Napjaink vegyipari technológiáit lassan kivétel nélkül számítógép felügyelte irányítórendszer segít üzemeltetni. Ezen rendszerek azonban már nem csupán az alapvető szabályozási feladatokat látják el, hanem feladatukká vált a működés közben keletkezett folyamatadatok tárolása, naplózása is. Ennek következtében az üzemmenet során lenyűgöző mennyiségű adatot - ezen keresztül információt - rögzítenek.

Ennek köszönhetően a technológiafejlesztési módszerek tárháza is bővülni kezd, hiszen az eddig használt tapasztalati eszközök mellé felsorakoznak a matematikai modellezés és optimalizálás metódusai. A statisztikai eszközök - pl.: SPC (Statistical Process Contol), PCA (Principal Comonent Analysis - Főkomponens-elemzés) alapú hibadetektálás - mellett a fizikai-kémiai összefüggéseket explicit módon felhasználó alkalmazások széles skálája biztosítja a különböző mérnöki feladatok hatékony megoldását, a technológia fejlesztését. Ezen technikákban pedig egy mindenképpen közös: az eredményes felhasználás záloga az éppen gyűjtött illetve már tárolt adatokban rejlő információ kiaknázása úgy, hogy a folyamat bemeneti és kimeneti adatai közti összefüggéseket is vizsgáljuk, azaz modelleket készítünk.

Minden modell alapú fejlesztés első lépése, a technológia megfelelő megismerése után, a kívánt modellezési út meghatározása és a modell alapvető struktúrájának kialakítása. Ezen struktrúra megalkotása az adott cél függvénye: az APC (Advanced Process Control) eszközök általában lineáris modelleket, míg pl. az OTS (Operator Training System) eszközök nemlineáris, fehér

doboz modelleket használnak. Minden modellezési feladat következő lépése a modellparaméterek meghatározása, amihez ki kell válogatni a megfelelő bemenet-kimenet folyamatadatokat. Ez nehéz feladat, hiszen ki kell gyomlálni az adott működési zavarok időszakait, meg kell határozni a megfelelően lineáris vagy nemlineáris működési periódusok adathalmazait. Ezen adathalmazok végigkísérik a modellezés lépéseit az identifikálástól kezdve a validálásig, de fontos szerepet játszthatnak végső alkalmazás gazdaságossági vizsgálatának eredményében is.

A dolgozat célja olyan innovatív eszközök bemutatása, amelyek a gyűjtött folyamatadatok szegmentálásával segítik a működési zavarok időszakainak elkülönítését, detektálják a bemenet-kimenet adatok közti lineáris korreláció megváltozását, illetve támogaták a paraméteridentifikálás során szükséges informatív adathalmazok szegregációját, emellett a folyamatadatok elemzésén alapuló, gazdasági eredmény elemzését figyelembe vevő technológiafejlesztést.

A kereskedelmi forgalomban lévő hibadatektálási metódusok jó része PCA alapú.Első lépésként kiválasztanak egy garantáltan hibamentes működési periódust, aminek adatait felhasználva kialakítják a PCA modellt. A továbbiakban összevetik a PCA alapú eszköz által predikált illetve a folyamatból gyűjtött kimeneti adatokat és a kettő különbsége alapján döntik el, hogy hiba bekövetkezett-e vagy sem. A dolgozatban bemutatott PCA alapú idősorszegmentálás során a cél kettős lehet: (i) hibadetektálás illetve (ii) a bement-kimeneti adatok közti lineáris kapcsolat megváltozása. Ez előbbi haszna nyilvánvaló, míg az utóbbi segít azon időszakaszok elkülönítésben, ahol a bement-kimeneti adatok közt lineáris kapcsolat áll fenn. Ezen információ például az Advanced Process Control (APC) alkalmazásokban eredményesen felhasználható. Köszönhetően annak, hogy mind on-line és off-line idősorszegmentálási módszert bemutattuk, így az alkalmazási lehetőségek tárháza is bővül, hiszen mind hisztorikus, mind on-line adatok is elemezhetővé válnak.

Megfelelő hibamentes adatok birtokában a modellek struktúrájának megalkotása utan az egyetlen feladattá a modellparaméterek becslésehez szükséges információgazdag időszakaszok elkülönítése válik. Ehhez a kísérlettervezéses Optimal Experiment Design (OED) eszközök tárházából kölcsönzött Fisher-információs mátrixot hívtuk segítségül. Ezen mátrix tartalmazza az adott bemeneti szekvencia mellett a modellkimenet paraméterérzékenységét, azaz a kimeneti v'altozók paraméterek szerinti parciális differenciálhányadosait. Ezt integrálva a klasszikus idősorszegmentálási eljárásaba egy új, innovatív idősorszegmentálási eszközt mutattam be, amely segítségével adott, előre meghatározott modelstruktúra mellett a paramétermeghatározás szempontjából információgazdag szegmensek

elkülöníthetők és az információtartalom az OED-ból ismert E és D kritériumok segítségével mérhetővé válik.

Mindezek

mellett a kísérletes optimalizálás eszközei is hatékonyan használhatók az adatalapú tecnológifejlesztés során a gazdasági hatékonyságnövelés egyik eszközeként. Ezen a területen leginkább elterjedté az APC alkalmazások váltak, amik alapját a modell prediktív szabályozók (Model Predicitve Controller - MPC) képezik. Az MPC-k hatásosságát - a modellek predikciós képessége mellett - nagyban befolyásolja a szabályozó paramétereinek megfelelő beállítása. Ez jelentős hatással van a gazdasági jellegű célfüggvénnyel mérhető teljesítményre, ami a folyamatadatok ismeretében számítható. A nehézséget általában az okozza, hogy az MPC-k hangolóparaméterei közt vannak olyanok, amelyek csak egész értékeket vehetnek fel. A gazdasági jellegű célfüggvényt a kísérletes optimalizálás eszközeivel integrálva (szimplex módszer, teljes faktoros tervek) egy hatékony keretrendszert mutattam be, amely a hangolóparaméterek változtatásával szisztematikusan halad a gazdasági optimum irányába a technológiai korlátokat és szűk keresztmetszeteket is figyelembe véve. Abban az esetben, ha a szabályozott objektum modellje rendelkezésre állna, úgy már a tervezés szakaszában meghatározhatók az optimális szabályozóparaméterek. Ha azonban ez nem lehetséges, úgy az iteratív tanuló szabályozási sémát használva, lépésről-lépésre haladhatunk a gazdasági optimum felé gyártási cikulsról-gyárási ciklusra.

6.2. Új tudományos eredmények

1. **Kimutattam, hogy a működő technológiák folyamatadatait többváltozós idősorként kezelve a dinamikus főkomponenselemzés és rekurzív kovariancia mátrix számítás eszközeit alapul véve olyan on-line és off-line módon alkalmazható idősorszegmentáló eszközt fejlesztettem, amely alkalmas az esetleges zavarások, meghibásodások, illetve egyéb ok okozta működési tartomány változások detektálására és a homogén működési tartományok elkülönítésére.**
   (Kapcsolódó publikációk: 4, 7, 12)

   A működő technológiák egyik fejlesztési lehetősége a homogén működési tartományok feltárása és a különböző működési tartományok elkülönítése és

az ezeket okozó zavarások és meghibásodások időpontjainak meghatározása. Ennek célja a zavarás, meghibásodás okának kiderítése vagy csupán az adott bemenet-kimenet adatok közti linearitás változásának detektálása. Ez leginkább a lináris modell prediktív szabályozók alkalmazásakor nyer igazán értelmet, hiszen az alkalmazott lineáris modell adott működési tartományban validitását vagy elvesztését jelzi.

A többváltozós adatsorokban a változók közti kapcsolatok vizsgálatára elterjedt módszer a főkomponens elemzés (Principal Component Analysis - PCA). Mivel ez a statisztikai módszer nincs tekintettel a folyamatadatok időbeliségére így ezek figyelembe vételére a dinamikus főkomponens elemzést (dynamic PCA - dPCA) alkalmaztam, ahol az eredeti adatmátrixot kiegészítik az előző időpillanatokban mért folyamatadatok értékeivel. Statisztikai módszer lévén a kovariancia mátrixok számításához nagy számú adat szükséges, ami a zavarások időpontjának pontos detektálásánál hátrányt jelent, így felbontás növelésére felmerült az igény arra, hogy minden mintavételezéskor rendelkezésre álljon egy kovariancia mátrix. Ennek számítására a rekurzív megközelítést használtam, ahol az előző mintavételezési időbeli kovariancia mátrix és az aktuális mintavételezési időpontban gyűjtött folyamatadatok segítségével kiszámítható az aktuális mintavételezési időpontbeli kovariancia mátrix. Ezzel minden mintavételezési időben rendelkezésre áll egy kovariancia mátrix. A rekurzív számítás egyik paramétere a felejtési tényező, ami az előző mintavételezési időben kalkulált kovariancia mátrix és az aktuális folyamatadatok súlyozására szolgál. A kovariancia mátrixok gyors és hatékony adaptációjához a Fortescue féle változó értékű felejtési tényezőt használtam. A kovariancia mátrixok alkotta idősorban a mátrixok hasonlóságát a Krzanowski-féle hasonlóságmértékkel határoztam meg, ami gyakorlatilag a dPCA modellek által reprezentált hipersíkok bezárt szögének cosinusának kiszámítása.

Ezeket az eszközöket a bottom-up (lentről felfele történő) és a sliding window (csúszóablakos) idősorszegmentálási technikákba integrálva off-line illetve on-line idősorszegmentálási algoritmust fejlesztettem, amik segítségével naplózott illetve valós idejű folyamatadok alapján azonnal detektálható az esetleges zavarás, meghibásodás. A kialakított keretrendszer működését egy benchmark példán, a Tennesse Eastman problémán is megvizsgáltam.

2. Az optimális kísérletek tervezésekor használt Fisher információs mátrix és az ezen alapuló D és E kritérium segítségével olyan alkalmazást készítettem, amely a működő technológiákban gyűjtött bemenet-kimenet idősorokat olyan homogén idősorszegmensekre bontja, amelyek egy folyamatmodell adott paraméterkészletének identifikálásakor eltérő információtartalommal bírnak.
(Kapcsolódó publikációk: 1, 5, 10, 11)

A folyamatmodelleken alapuló eszközök alkalmazás'anak elterjedésével egyre nő az igény a megbízhatóan működő, a folyamatot jól reprezentáló modell elkészítésére. Ezek egyik sarokköve a modellparaméterek pontos meghatározása. A paraméteridentifikáció egy kulcspontja pedig azon bemenet-kimenet adasorok elkülönítése és biztosítása, amelyekből a modellparaméterek pontosan meghatározhatók, azaz adatsorok információban gazdagok legyenek. Ezen adatok biztosításának egyik módja a megfelelő kísérletek tervezése és kivitelezése, ami idő-, erőforrás- és tőkeigényes, másik módja pedig a naplózott folymatadatokból való információgazdag idősor részletek elkülönítése.

Az optimális kísérletek tervezésekor használt eszközök segítségével egy adatsor információtartalma meghatározható. Ezek alapja a Fisher információs mátrix, amely a modellkimenet modellparaméterek szerinti parciális deriváltjaiból (érzékenység egyenletekből) számítható adott bemeneti adatok mellett, így a bemeneti adatok információtartalmára vonatkozóan kapunk egyfajta kovariancia mátrixot. A Fisher információs mátrixból pedig a D és E kritériumok segítségével egy skalár mérőszámmal meghatározható az információtartalom nagysága.

A folyamatból gyűjtött hisztorikus adatokat felhasználva minden mintavételezési időpontban kiszámítottam az érzékenység egyenletek alapján a modellkimenet paraméterváltozásra történő érzékenységét, egy parciális differenciálhányadosokból álló idősort kaptam. Az így kapott idősor alapján generálható egy Fisher információs mátrix idősor, amik alapján a D és E kritériumok értéke kiszámítható.

A Fisher információs mátrix azonban az adott paraméterkészlet esetén nem csupán a paraméterkészletre vonatkozó információtartalom nagyságára, hanem a paraméterkészlet terében az irányára vonatkozó információkat is

tartalmaz. Ezen információt a PCA kovariancia mátrixához hasonlóan a sajátvektorok hordozzák és bemutattam, hogy így lehetővé válik az ezeken alapuló idősorszegmentálás is, ahol a Fisher mátrixok hasonlóságát a Krzanowski hasonlósági mértékkel határoztam meg. Ezt integráltam a bottom-up idősorszegmentálási technikába, így egy off-line idősorszegmentálási eszközt készítettem, aminek segítségével az információtartalom iránynak megváltozása az információs térben is detektálható.

3. **Olyan keretrendszert mutattam be, ahol a klasszikus kísérletes optimalizálási technikák hatékonyan és célirányosan alkalmazhatók a működő technológiák modell prediktív alapú irányírási rendszerének gazdasági jellegű tervezésében, optimalizálásában és működtetésében.**
(Kapcsolódó publikációk: 2, 3, 6, 8, 9, 13)

A folyamatirányító rendszerek fejlődésével egyidőben felmerül a kérdés, hogy a modern irányítási rendszerek alkalmazásával nyert haszon hogyan számszerűsíthető és ez a haszon miként maximalizálható. Mivel a legújabb irányítási rendszerekben modell prediktív szabályozók alkalmazása széles körben elterjedt, így adódik az igény azon hangolóparaméterek meghatározására, amelyek segítségével a legkedvezőbb gazdasági teljesítmény érhető el. Különösen fontos ez a termék- és munkapontváltások esetén, ahol a cél pl. a legkevesebb selejtanyag előállítása.

Az általam kidolgozott eljárás alapja egy gazdasági jellegű célfüggvény, amely tartalmazza az előállított termék (mind a specifikációt teljesítő és nem teljesítő termék) értékesítési árát, az alapanyag beszerzési árát és a gyártás közben felmerülő egyéb költségeket, amelyek alapján a gazdasági haszon maximalizálható, az adott hangolóparaméterek által biztosított gazdasági teljesítmény mérhető. Bemutattam, hogy a szimplex módszer alkalmazásával a modell prediktív szabályozók hangolóparamétereinek (predikciós horizontok és büntető tagok) köszönhetően a vegyes-egész értékű optimalizálási feladat jól kezelhető, figyelembe véve a rendszerben működtetése során felmerülő operatív korlátokat is.

(a) Amennyiben a szabályozott

objektum matematikai modellje rendelkezésre áll és megalkotható a szabályozott rendszer - irányító rendszer matematikai modellje úgy szimulációval meghatározható, hogy a termék- és munkapontváltások az adott hangolóparaméterek hatására milyen gazdasági eredménnyel realizálódnának. Ezt a szimulátort a szimlex módszer keretrendszerébe helyeztem amivel meghatározható a legnagyobb gazdasági gazdasági haszonnal járó hangolóparaméter kombináció.

(b) Amennyiben a szabályozott objektum matematikai modellje nem áll rendelkezésre úgy az iterative learning control sémába behelyettesítve az operátorok a kíséletes optimalizálás eszközeit és a gazdasági célfüggvényt használva a hangolóparaméterek értékeit kiszámolhatják és manuálisan is beállíthatják.

6.3. További kutatási lehetőségek

A dolgozatban bemutatott algoritmusok és a kapcsolódó eredmények számos új érdekes továbblépési lehetőséget vetnek fel.

- A jelenleg alkalmazott mérnöki módszerek támogatására a dPCA és a Fisher információs mátrix alapú szegmentálási algoritmusok egymásba integrálása hatalmas segítséget jelenthet pl. Advanced Process Control (APC) projektek kivitelezése kapcsán. Ez azt jelentené, hogy az irányitórendszerből (Distributed Control System - DCS) gyűjtött adatok először a linearitás szempontjából szegmentálhatók és az így keletkezett szegmensekben pedig a Fisher féle információs mátrix alapján elkülöníthetők az eltérő információtartalmú periódusok. Ezek felhasználóbarát implentálása az APC projektek egyik időigényes lépését, a lépésteszt elést nagyban rövidíthetné. A lépésteszt során a működő technológia bemeneti változóin egysegugrást gerjesztést alkalmaznak, hogy végül a bemenet-kimenet modellek identifikálhatók legyenek.

- Emellett érdekes kutatási terület lehet az Independent Component Analysis (ICA) és ennek dinamikus változatának használata a hibadetektálás során, amikor is a PCA modell helyett az ICA eszközök segítségével teremthető kapcsolat a bemenet-kimenet adatok közt.

- A dinamikus PCA alapú szegmentálás egyik hátrányaként említhető, hogy a linearitás megváltozásának csak az időpontját adja meg, míg a kiváltó okot nem tudja meghatározni. Hasonlóan az Artificial Immune System (AIS) alapú technikákhoz, hasznos lehet megvizsgálni, hogy a PCA alapú szegmentálási algoritmus kiterjeszthető-e a hiba valószínűsített okának meghatározásához is a felügyelt tanítási (supervised learning) technikák segítségével.

- Az idősorszegmentálási technikák mellett a szimplex alapú szabályozóhangolási metódus továbbfejlesztése is érdekes terület lehet, hiszen a korábbi működés során gyűjtött hisztorikus adatok felhasználása a pl. módszer inicializálásakor nem került részletes vizsgálatra.

Appendix A

Tables in the thesis

Table A.1: Process disturbances for the Tennessee Eastmen Process

Case	Disturbance	Type
1	A/C feed ratio, B composition constant	Step
2	B composition ratio, A/C feed constant	Step
3	D feed temperature	Step
4	Reactor cooling water inlet temperature	Step
5	Condenser cooling water inlet temperature	Step
6	A feed loss	Step
7	C header pressure loss - reduced availability	Step
8	A, B, C feed composition	Random variation
9	D feed temperature	Random variation
10	C feed temperature	Random variation
11	Reactor cooling water inlet temperature	Random variation
12	Condenser cooling water inlet temperature	Random variation
13	Reaction kinetics	Slow drift
14	Reactor cooling water valve	Sticking
15	Condenser cooling water valve	Sticking
16-20	Unknown	Unknown

Table A.2: Process variables used for dynamic PCA

Output variables	A feed D feed E feed A and C feed Recycle flow Reactor feed rate Reactor temperature Purge rate Product separator temperature Product separator pressure Product separator underflow Stripper pressure Stripper temperature Stripper steam flow Reactor cooling water outlet temperature Separator cooling water outlet temperature
Manipulated variables	D Feed E Feed A Feed A+C Feed Purge Valve Separator Valve Stripper Valve Reactor Coolant Condenser Coolant

Table A.3: Result of identification scenarios of determining all kinetic parameters

Scenario	Segment	$\frac{cost}{1000 samples}$	Type of parameter	r=p	r=fm	r=I	r=tc	r=td
Original	-	-	k_r	$1.77 \cdot 10^9$	$1.0067 \cdot 10^{15}$	$3.792 \cdot 10^{18}$	$3.8223 \cdot 10^{10}$	$3.1457 \cdot 10^{11}$
			E_r	$1.8283 \cdot 10^4$	$7.4478 \cdot 10^4$	$1.2877 \cdot 10^5$	$2.9442 \cdot 10^3$	$2.9442 \cdot 10^3$
Best	5^{th}	$7.71 \cdot 10^7$	k_r	$2.13 \cdot 10^9$	$1.75 \cdot 10^{15}$	$2.33 \cdot 10^{18}$	$2.53 \cdot 10^{10}$	$1.57 \cdot 10^{11}$
			E_r	$2.07 \cdot 10^4$	$8.16 \cdot 10^4$	$1.25 \cdot 10^5$	$3.95 \cdot 10^3$	$2.33 \cdot 10^3$
Worst	1^{st}	$3.38 \cdot 10^8$	k_r	$2.85 \cdot 10^9$	$9.09 \cdot 10^{14}$	$7.58 \cdot 10^{18}$	$4.42 \cdot 10^{10}$	$2.77 \cdot 10^{11}$
			E_r	$2.18 \cdot 10^4$	$7.97 \cdot 10^4$	$1.29 \cdot 10^5$	$2.11 \cdot 10^3$	$2.36 \cdot 10^3$

Table A.4: Result of identification scenarios of determining exponential parameters

Scenario	Segment	$\frac{cost}{1000 samples}$	Type of parameter	r=p	r=fm	r=I	r=tc	r=td
Best	5^{th}	$4.0538 \cdot 10^3$	E_r	$1.8626 \cdot 10^4$	$7.4832 \cdot 10^4$	$1.2917 \cdot 10^5$	$5.7232 \cdot 10^3$	$3.4302 \cdot 10^3$
Worst	2^{nd}	$1.9573 \cdot 10^6$	E_r	$1.9978 \cdot 10^4$	$7.7165 \cdot 10^4$	$1.2775 \cdot 10^5$	$4.4439 \cdot 10^3$	$5.2659 \cdot 10^3$

Table A.5: Result of identification scenarios of determining preexponential parameters

Scenario	Segment	$\frac{cost}{1000samples}$	Type of parameter	r=p	r=fm	r=I	r=tc	r=td
Best	5^{th}	165	k_r	$1.6281 \cdot 10^{9}$	$9.2605 \cdot 10^{14}$	$4.3787 \cdot 10^{18}$	$6.6971 \cdot 10^{10}$	$2.4946 \cdot 10^{11}$
Worst	2^{nd}	$1.6794 \cdot 10^{7}$	k_r	$2.1630 \cdot 10^{9}$	$1.6956 \cdot 10^{15}$	$2.3419 \cdot 10^{18}$	$2.6473 \cdot 10^{10}$	$1.5729 \cdot 10^{11}$

Appendix B

Publications related to theses

Articles in International Journals

1. L. Dobos, Z. Bankó, J. Abonyi. Qualifying and segmentation of historical process data using optimal experiment design techniques for supporting parameter estimation. *Acta Electrotechnica et Informatica*, Vol. 10, pp. 28-32, 2010.

2. L. Dobos, J. Abonyi. Controller Tuning of District Heating Networks using Experiment Design Techniques. *Chemical Engineering Transactions*, Vol. 21, DOI: 10.3303/CET1021239, 2010.

3. L. Dobos, J. Abonyi. Controller tuning of district heating networks using experiment design techniques. *ENERGY*, DOI: 10.1016/j.energy.2011.04.014

4. L. Dobos, J. Abonyi. On-line detection of homogeneous operation ranges by dynamic principal component analysis based time-series segmentation. *Chemical Engineering Science*, Vol 75, pp 96-105, 2012.

5. L. Dobos, J. Abonyi. Fisher information matrix based time-series segmentation of multivariate streaming data for supporting model identification. *Chemical Engineering Science*, Vol 101, pp. 99-108, 2013.

Articles in Hungarian Journals

6. L. Dobos, J. Jäschke, J. Abonyi, S. Skogestad. Dynamic model and control of heat exchanger networks for district heating. *Hungarian Journal of Industrial Chemistry*, Vol. 37(1), pp 37-49, 2009.

Refereed Presentations

7. L. Dobos, Z. Bankó, J. Abonyi. Optimal experiment design integrated with time-series segmentation. *Proceedings of the 8th International Symposium on Applied Machine Intelligence and Informatics*, pp 206-210, 2010.

8. L. Dobos, J. Abonyi. Application of non-linear dynamic optimization in advanced process control of product grade transitions of polymerization processes. *20th European Symposium on Computer Aided Process Engineering*, pp 559-564, 2010.

9. L. Dobos, J. Abonyi. Controller tuning of district heating networks using experiment design techniques. *19th International Congress of Chemical and Process Engineering CHISA 2010*, PRES 2010, 2010.

10. L. Dobos, J. Abonyi. Fisher information based time-series segmentation of streaming process data for monitoring and supporting on-line parameter estimation in energy systems. *21th European Symposium on Computer Aided Process Engineering*, pp 1844-1849, 2011.

11. L. Dobos, Z. Bankó, J. Abonyi. Segmentation based optimal experiment design. *2009 Proceedings of the 10th International Symposium of Hungarian Researchers*, pp 279-289, 2009.

12. L. Dobos, J. Abonyi. Application of on-line multivariate time-series segmentation for process monitoring and control. *11th International PhD Workshop on Systems and Control*, 2010.

13. L. Dobos, J. Abonyi. A novel economic oriented performance measure and tuning method for model predictive controllers. *Distributed Control Systems 16, International Scientific Workshop*, 2011.

Bibliography

[1] A. T. M. Coxon. *Collection and Analysis (Quantitative Applications in the Social Sciences)*. Sage Publications, 1993.

[2] B. Balasko and J. Abonyi. What happens to process data in chemical industry? from to applications - an overview. *Hungarian Journal of Industrial Chemistry*, 35:75–84, 2007.

[3] S. Kivikunnas. Overview of process trend analysis methods and applications. In *ERUDIT Workshop on Applications in Pulp and Paper Industry*, 1998. CD-ROM.

[4] S. Wold. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*.

[5] Manabu Kano, Shinji Hasebe, Iori Hashimoto, and Hiromu Ohno. A new multivariate statistical process monitoring method using principal component analysis. *Computers and Chemical Engineering*, 25:1103–1113, 2001.

[6] Li Gang, Qin Si-Zhao, Ji Yin-Dong, and Zhou Dong-Hua. Total pls based contribution plots for fault diagnosis. *Acta Automatica Sinica*, 35(6):1103–1113, 2009.

[7] Y.M. Sebzalli and X.Z. Wang. Knowledge discovery from process operational data using pca and fuzzy clustering. *Engineering Applications of ArtiiňĄcial Intelligence*, 14:607–616, 2001.

[8] Z. Ge, C. Yang, and Z. Song. Improved kernel pca-based monitoring approach for nonlinear processes. *Chemical Engineering Science*, 64:2245–2255, 2009.

[9] R. Luo, M. Mistra, and D. M. Himmelblau. Sensor fault detection via multiscale analysis and dynamic pca. *Industrial and Chemical Engineering Research*.

[10] R. Shao, F. Jia, E.B. Martin, and A.J. Morris. Wavelets and non-linear principal components analysis for process monitoring. *Control Engineering Practice*, 7:865–879, 1999.

[11] Yingwei Zhang and Chi Ma. Fault diagnosis of nonlinear processes using multiscale kpca and multiscale kpls. *Chemical Engineering Science*, 66:64–72, 2011.

[12] M. R. Maurya, R. Rengaswamy, and V. Venkatasubramanian. Fault diagnosis by qualitative trend analysis of principal components. *Chemical Engineering Research and Design*, 83:1122–1132, 2005.

[13] J. Chen and W. J. Chang. Applying wavelet-based hidden markov tree to enhancing performance of process monitoring. *Chemical Engineering Science*, 60:5129–5143, 2005.

[14] S. X. Wu and W. Banzhaf. The use of computation intelligence in intrusion detection systems: A review. *Applied Soft Computing*, 10:1–35, 2010.

[15] C. A. Laurentys, R. M. Palhares, and W. M. Caminhas. A novel artificial immune system for fault behavior detection. *Expert Systems with Applications*, 38:6957–6966, 2011.

[16] Y. Dai and J. Zhao. Fauld disagnosis of batch chamical processes using dynamic time warping (dtw)-based artificial immune system. *Industrial and Engineering Chemistry Research*, 50:4534–4544, 2011.

[17] M. Last, Y. Klein, and A. Kandel. Knowledge discovery in time series databases. *IEEE Transactions on Systems, Man, and Cybernetics*, 31(1):160–169, 2000.

[18] T. C. Fu. A review on time series data mining. *Engineering Applications of Artificial Intelligence*, 24:164–181, 2011.

[19] M. E. Tipping and C. M. Bishop. Mixtures of probabilistic principal components analysis. *Neural Computation*, 11:443–482, 1999.

[20] M. Bin Shams, H. Budman, and T. Duever. Finding a trade-off between observability and economics in the fault detection of chemical processes. *Computers and Chemical Engineering*, 2010. In press.

[21] W. Li, H. Henry Yue, S. Valle-Cervantes, and S. Joe Qin. Recursive pca for adaptive process monitoring. *Journal of Process Control*, 10:471–486, 2000.

[22] T.L. Fortescue, L.S. Kershenbaum, and B.E. Ydstie. Implementation of self-tuning regulators with variable forgetting factors. *Automatica*, 17:831–835, 1981.

[23] W. Ku, R.H. Storer, and C. Georgakis. Disturbance detection and isolation by dynamic principle component analysis. *Chemometrics and Intelligent Laboratory Systems*, 30:179–196, 1995.

[24] S. W. Choi and I. B. Lee. Nonlinear dynamic process monitoring based on dynamic kernel pca. *Chemical Engineering Science*, 59:5897–5908, 2004.

[25] E. Keogh, S. Chu, D. Hart, and H. Pazzani. An online algorithm for segmenting time series. *IEEE International Conference on Data Mining*, 2001. URL: http://citeseer.nj.nec.com/keogh01online.html.

[26] E. Keogh, S. Chu, D. Hart, and M.J. Pazzani. An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback. *4th Int. Conf. on KDD.*, pages 239–243, 1998.

[27] W.J. Krzanowsky. Between group comparison of principal components. *J. Amer. Stat. Assoc.*, pages 703–707, 1979.

[28] A. Singhal and D.E. Seborg. Matching patterns from historical data using PCA and distance similarity factors. *Proceedings of the American Control Conference*, pages 1759–1764, 2001.

[29] Y. Cao and P. Saha. Improved branch and bound method for control structure screening. *Chemical Engineering Science*, 60:1555–1564, 2005.

[30] M.A. Bin Shams, H.M. Budman, and T.A. Duever. Fault detection, identiiňĄcation and diagnosis using cusum based pca. *Chemical Engineering Science*, doi:10.1016/j.ces.2011.05.028, 2011.

[31] J. M. Lee, C.K. Yoob, and I. B. Lee. Statistical monitoring of dynamic processes based on dynamic independent component analysis. *Chemical Engineering Science*, 59:2995–3006, 2004.

[32] R. Srinivasan and M. S. Qian. Online fault diagnosis and state identiïňĄcation during process transitions using dynamic locus analysis. *Chemical Engineering Science*, 61:6109–6132, 2006.

[33] K. Bernaerts, R. D., S. Kooyman, K. J. Versyck, and J. F. van Impe. Optimal temperature design for estimation of the square root model parameters: parameter accuracy and model validity restrictions. *Int. Jour. Of Food Microbiology*, 73:147–157, 2002.

[34] C.E. Shannon. A mathematical theory of communication. *Bell System Tehcnical Journal*.

[35] Gaia Franceschini and Sandro Macchietto. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63:4846–4872, 2008.

[36] S.P. Asprey and S. Macchietto. Designing robust optimal dynamic experiments. *Journal of Process Control*, 12:545–556, 2002.

[37] C. R. Rojas, J. S.Welsh, G. C. Goodwin, and A. Feuer. Robust optimal experiment design for system identification. *Automatica*, 43:993–1008, 2007.

[38] T. Turanyi. Sensitivity analysis of complex kinetic systems. tools and applications. *Journal of Mathematical Chemistry*, 5:203–248, 1990.

[39] K. Bernaerts, K. P. M. Gysemans, T. N. Minh, and J. F. van Impe. Optimal experiment design for cardinal values estimation: guidelines for data collection. *International Journal of Food Microbiology*, 100:153–165, 2005.

[40] D. Telen, F. Logist, E. Van Derlinden, I. Tack, and J. Van Impe. Optimal experimentdesign for dynamic bioprocesses: A multi-objective approach. *Chemical Engineering Science*, 78:82–97, 2012.

[41] Andre L. Alberton, Marcio Schwaab, Marcos Wandir Nery Lobao, and Jose Carlos Pinto. Experimental design for the joint model discrimination and precise parameter estimation through information measures. *Chemical Engineering Science*, 66:1940–1952, 2012.

[42] Brecht M.R. Donckels, Dirk J.W. De Pauw, Peter A. Vanrolleghem, and Bernard De Baets. An ideal point method for the design of compromise experiments to simultaneously estimate the parameters of rival mathematical models. *Chemical Engineering Science*, 65:1705–1719, 2010.

[43] Andre L. Alberton, Marcio Schwaab, Marcos Wandir Nery Lobao, and Jose Carlos Pinto. Design of experiments for discrimination of rival models based on the expected number of eliminated models. *Chemical Engineering Science*, 75:120–131, 2012.

[44] A. Ashoori, B. Moshiri, A. Khaki-Segigh, and M.R. Bakhtiari. Optimal control of a nonlinear fed-batch fermentation process using model predictive approach. *Journal of Process Control*, 19:1162–1173, 2009.

[45] D.H. Owens and J. Hatonen. Iterative learning control ï£¡ an optimization paradigm. *Annual Reviews in Control*, 29:57–70, 2005.

[46] Tao Liu, Furong Gao, and Youqing Wang. Imc-based iterative learning control for batch processes with uncertain time delay. *Journal of Process Control*, 20:173–180, 2010.

[47] A. Flores-Tlacuahuac and A. Biegler. Integrated control and process design during polymer grade transition operations. *Computers and Chemical Engineering*, 32:2823–2837, 2008.

[48] S. BenAmor, F. J. Doyle, and R. McFarlane. Polymer grade transition control using advanced real-time optimization software. *Journal of Process Control*, 14:349–364, 2004.

[49] K.B. McAuley and J.F. MacGregor. On-line inference of polymer properties in an industrial polyethylene reactor. *AIChe Journal*, 31:825–835, 1991.

[50] Y. Farzaneh and A.A. Tootoonchi. A novel data reduction method for takagi-sugeno fuzzy system design based on statistical design of experiments. *Applied Soft Computing*, 9:1367–1376, 2009.

[51] K. Vasko and H.T.T. Toivonen. Estimating the number of segments in time series data using permutation tests. *IEEE International Conference on Data Mining*, pages 466–473, 2002.

[52] J. Himberg, K. Korpiaho, H. Mannila, J. Tikanmaki, and H.T. Toivonen. Time-series segmentation for context recognition in mobile devices. *IEEE International Conference on Data Mining (ICDMÃ■01), San Jose, California*, pages 466–473, 2001.

[53] B. Huang. Process identification based on last principal component analysis. *Journal of Process Control*, 11:19–33, 2001.

[54] J. Wang and S. J. Qin. A new subspace identiïňĄcation approach based on principal component analysis. *Journal of Process Control*, 12:841–855, 2002.

[55] S. Lane, E.B. Martin, A.J. Morris, and P. Gover. Application of exponentially weighted principal component analysis for the monitoring of a polymer film manufacturing process. *Transactions of the Institute of Measurement and Control*, 25:17–35, 2003.

[56] S.J. Wieda. Multivariate statistical process control- recent results and directions for future research. *Statistica Neerlandica*, 48:147–168, 1994.

[57] N.D. Tracy, J.C. Young, and R.L. Mason. Multivariate control charts for individual observations. *Journal of Quality Technology*, 24:88–95, 1992.

[58] J.J. Downs and E.F. Vogel. A plant-wide industrial process control problem. *Computers and Chemical Engineering*, 17:245–255, 1993.

[59] N.L. Ricker. Decentralized control of the tennessee eastman challenge process. *Journal of Process Control*, 6:205–221, 1996.

[60] Janos Abonyi, Balazs Feil, Sandor Nemeth, and Peter Arva. Modified gath and ÂĂÂŞgeva clustering for fuzzy segmentation of multivariate time-series. *Fuzzy Sets and Systems*, 149(1):39 – 56, 2005.

[61] L. Dobos and J. Abonyi. On-line detection of homogeneous operation ranges by dynamic principal component analysis based time-series segmentation. *Chemical Engineering Science*, 75:96–105, 2012.

[62] G.C. Goodwin. Optimal input signals for nonlinear-system identification. *Proceedings of the Institution of Electrical Engineers*, 118:922–926, 1971.

[63] R. Mehra. Optimal input signals for parameter estimation in dynamic systems - survey and new results. *IEEE Transactions on Automatic Control*, 19:753–768, 1974.

[64] Z. D. Yuan and L. Ljung. Black-box identification of multivariable transfer functionsï£¡asymptotic properties and optimal input design. *International Journal of Control*, 40:233–256, 1984.

[65] M. Morari and J. H. Leeb. Model predictive control: past, present and future. *Computers and Chemical Engineering*, 23:667ï£¡682, 1999.

[66] R. Hildebrand and M. Gevers. Identification for control: Optimal input design with respect to a worst-case ν-gap cost function. *SIAM J. Control Optim.*, 41(5):1586ï£¡1608, 2002.

[67] A. Silva-Beard and A. Flores-Tlacuahuac. Effect of process design/operation on steady-state operability of a methyl-metacrylate polymerization reactor. *Ind. Eng. Chem. Res.*, 38:4790–4804, 1999.

[68] M.F. Torchio, G. Genon, and M. Poggio A. Poggio and. Merging of energy and environmental analyses for district heating systems. *Energy*, 34:220–227, 2009.

[69] S.Y. Choi, K.Y. Yoo, J.B. Lee, C.B. Shin, and M.J.Park. Mathematical modeling and control of thermal plant in the district heating system of korea. *Applied Thermal Engineering*, 30:2067–2072, 2010.

[70] T. Savola and I. Keppo. Off-design simulation and mathematical modeling of small-scale chp plants at part loads. *Applied Thermal Engineering*, 25:1219–1232, 2005.

[71] T. Savola, T.M. Tveit, and C.J. Fogelholm. A minlp model including the pressure levels and multiperiods for chp process optimisation. *Applied Thermal Engineering*, 27:1857–1867, 2007.

[72] A. Molyneaux, G. Leyland, and D. Favrat. Environomic multi-objective optimisation of a district heating network considering centralized and decentralized heat pumps. *Energy*, 35:751–758, 2010.

[73] H. Lund, B. Möller, B.V. Mathiesen, and A. Dyrelund. The role of district heating in future renewable energy systems. *Energy*, 35:1381–1390, 2010.

[74] A. Benonysson, B. Bohm, and H.F. Ravn. Operational optimization in a district heating system. *Energy Conversion and Management*, 36:297–314, 1995.

[75] H.F. Ravn and J.M. Rygaard. Optimal scheduling of coproduction with a storage. *Engineering optimization*, 22:267–281, 1994.

[76] H. Zhao, J. Holst, and J. Arvastson. Optimal operation of coproduction with storage. *Energy*, 23:859–866, 1998.

[77] D. Knutsson, J. Sahlin, S. Werner, T. Ekvall, and E.O. Ahlgren. Heatspot-a simulation tool for national district heating analyses. *Energy*, 31:278–293, 2006.

[78] G. Sandou, S. Font, S. Tebanni, A. Hiret, and C. Mondon. Predictive control of a complex district heating network. In *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference, Seville, Spain, December 12-15*, pages 7372–7377, 2005.

[79] T.R. Tveit, T. Savola, A. Gebremedhin, and C.J. Fogelholm. Multi-period minlp model for optimising operation and structural changes to chp plants in district heating networks with long-term thermal storage. *Energy Conversion and Management*, 50:639–647, 2009.

[80] K.C.B. Steer, A. Wirth, and S.K. Halgamuge. Control period selection for improved operating performance in district heating networks. *Energy and Buildings*, doi:10.1016/j.enbuild.2010.10.029, 2010.

[81] H.T. Søgaard. *Stochastic systems with embedded parameter variations - Applications to district heating, PhD dissertation*. Technical University of Denmark, Institute of Mathematical Statistics and Operations Research, 1993.

[82] P. Pinson, T.S. Nielsen, H.Aa. Nielsen, N.K. Poulsen, and H. Madsen. Temperature prediction at critical points in district heating systems. *European Journal of Operational Research*, 194:163–176, 2009.

[83] A.H. Nielsen and H. Madsen. Modelling the heat consumption in district heating systems using a grey-box approach. *Energy and Buildings*, 38:63–71, 2006.

[84] S. Liu and G.P. Henze. Experimental analysis of simulated reinforcement learning control for active and passive building thermal storage inventory part 1. theoretical foundation. *Energy and Buildings*, 38:142–147, 2006.

[85] D. Konstantinidis, P. Varbanov, and J. Klemeš. Multi-parametric control and optimisation of a small scale chp. *Chemical Engineering Transactions*, 21:DOI: 10.3303/CET1021026, 2010.

[86] N. Yildirim, M. Toksoy, and G. Gokcen. Piping network design of geothermal district heating systems: Case study for a university campus. *Energy*, 35:3256–3262, 2010.

[87] L. Dobos, J. Jäschke, J. Abonyi, and S. Skogestad. Dynamic model and control of heat exchanger networks for district heating. *Hungarian Journal of Industrial Chemistry*, 37:37–49, 2009.

[88] S.S. Voutetakis, P. Seferlisa, S. Papadopoulou, and Y. Kyriakos. Model-based control of temperature and energy requirements in a fluidised furnace reactor. *Energy*, 31:2418–2427, 2006.

[89] D. Dougherty and D. Cooper. A practical multiple model adaptive strategy for multivariable model predictive control. *Control Engineering Practice*, 11:649–664, 2003.

[90] E.N. Pistikopoulos, P. Liu, and M.C. Georgiadis. Modelling and optimization issues of the energy systems of the future. *Chemical Engineering Transactions*, 21:DOI: 10.3303/CET1021001, 2010.

[91] K. Krüger, R. Franke, and M. Rode. Optimization ofboiler start-up using a nonlinear boiler model and hard constraints. *Energy*, 26:2239–2251, 2004.

[92] H. Deng, Z. Xu, and H.X. Li. A novel neural internal model control for multi-input multi-output nonlinear discrete-time processes. *Journal of Process Control*, 19:1392–1400, 2009.

[93] R. Shridhar and D.J. Cooper. A tuning strategy for unconstrained siso model predictive control. *Ind. Eng. Chem. Res.*, 36:729–746, 1997.

[94] N.M.M. Al-mhanna, H. Huebner, and R. Buchholz. Optimization of parameters growth conditions of yeast biomass during single cell protein production by using simplex method. *Chemical Engineering Transactions*, 21:DOI: 10.3303/CET1021080, 2010.

[95] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convegrence properities of the nelder-mead simplex method in low dimensions. *Siam Journal on Optimization*, 9:112–147, 1998.

[96] A. Kiraly, L. Dobos, and J. Abonyi. Economic oriented stochastic optimization in process control using taguchi's method. *Optimization and Engineering*, page Article in Press.

[97] C. Audet and J. J. E. Dennis. Mesh adaptive direct search algorithms for constrained optimization. *Siam Journal on Optimization*, 17:188–217, 2006.