Optimization and Online-Monitoring in industrial batch processes using Data-Mining methods

Christian Kühnert, Thomas Bernard

Fraunhofer-Institut für Informations- und Datenverarbeitung IITB
Fraunhoferstraße 1
Tel. (0721) 6091-9
{Christian.Kuehnert, Thomas.Bernard}@iitb.fraunhofer.de

Abstract

In general industrial batch processes are complex systems that have to be optimized due to several performance criteria. As control optimization based on the development of physical models is in many cases very time consuming and cost intensive or not even feasible an alternative approach consists in analyzing historical process data by means of computational intelligent methods. The aim is the identification of characteristic control patterns and classification of these patterns according to their contribution to process optimization. This paper presents a concept how the generated knowledge can be used for online monitoring production runs. The key idea is to derive a model from the historical process data which describes the impact of the most important features of the process variables (both manipulated and controlled variables) to a user-defined performance index (e.g. quality or losses of a production run). The set of relevant features is automatically found using an iterative algorithm based on Support Vector Machines. Further the model is used to calculate the optimal values of the relevant features for several segments. Based on the optimal values of the features an online monitoring of the process can be implemented. The proposed concept is applied to an industrial glass forming process.

1 Introduction

In general industrial batch processes are complex systems that have to be optimized due to several performance criteria. As control optimization based on the development of physical models is in many cases very time consuming and cost intensive or not even feasible an alternative approach consists in analyzing historical process data. Hereby the aim is the identification of characteristic control patterns and the classification of those patterns according to their contribution to process optimization. To reveal this information Data Mining techniques, which include machine learning theory, statistics and artificial intelligence, are increasingly used to find unknown and hidden correlations in the data [10], [14]. For optimization of complex systems these methods can be supplementary used besides physical models in order to determine unknown interconnections and important features in a process and thus can efficiently be applied for the optimization of process performance.

The key idea is to derive a model from the historical process data which describes the impact of the most important features of the process variables (both manipulated and controlled variables) to a user-defined performance index (e.g. quality or losses of a production run). A concept for a model generation and ranking of the relevant features has been already presented in [3], [4]. This paper is based on this concept, which is summarized in section 2. The new idea for an online monitoring is presented in section 3. Finally an application of the online monitoring to an industrial glass forming process is discussed in section 4.

2 Calculation of optimal features as a basis for online monitoring

This section summarizes a concept for model generation and ranking of relevant features [3], [4]. This model can then be used as the root for the online monitoring concept proposed in this paper. The model should describe the impact of the features to a user-defined performance index (see Fig. 1).

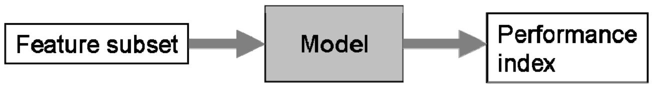

Fig. 1 Basis for the online monitoring is a subset of relevant features and a model which describes the impact of the features to a user-defined performance index. Both the relevant features as well as the model are derived from historical process data

For the model generation the work flow as shown in Fig. 2 has to be performed. In a first step raw process data has to follow some preprocessing steps (outlier removal, resampling and normalization of the data). This helps to increase the data quality coming from the process as well as reducing the amount of stored data [5]. Then new signals are generated (e.g. derivatives or combining different time series).

In a next step features have to be generated. This includes features coming from the complete time series (e.g. mean values over the whole production run) or features extracted from segments of the measured data. Simple examples of features are minima,

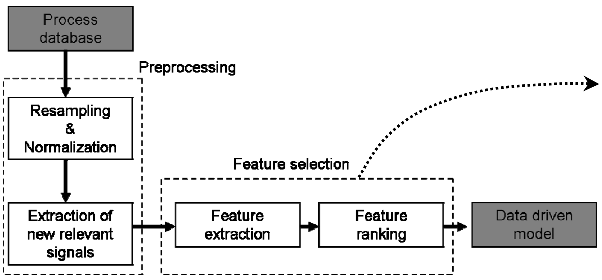

Fig. 2: Steps from measured process data to a data driven Model

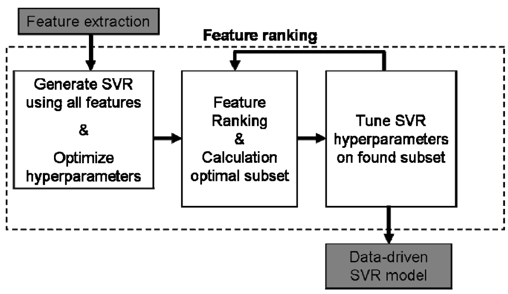

Fig. 3: Detailed work flow of feature ranking

maxima, mean values, trends of process variables or mathematically generated features like parameters from regression models. As not all of the extracted features have a meaningful relevance regarding the optimization of the process or tend to be strongly redundant to other features, a ranking procedure and the selection of an optimal feature subset has to be done. Feature ranking can be performed by discriminant or variance analysis as well nonlinear techniques [7], [8].

During all these steps it is necessary to include as much expert knowledge as possible, as the generation of suitable features is of crucial importance for the later derivation of a reliable model. The feature ranking and model building, as shown in detail in figure 3 is performed in several iterative steps. First a model based on Support Vector Machines using all features is generated. Then a first feature ranking is performed based on the support vectors coming from the model. Out of this ranking a first subset of features is calculated and the unimportant features are removed. Using the left features in a next step a new model is calculated and the model parameters are adapted accordingly. This procedure is repeated several times until the subset of features has converged to a final number. At last the resulting SVR model is calculated. This model will be used furthermore for the calculation of restrictions for online monitoring.

SVMs have the opportunity that a feature ranking can be integrated and besides the ranking procedure a nonlinear model can be generated by including kernel functions [11], [12]. As implementation least square Support Vector Machines are used [13].

By using the prior described concept a SVR model based on the main features of the process is derived. Afterwards this model is used to calculate the feature values corresponding to the wished value of the performance index. Usually the used performance index has to be maximized (e.g. efficiency) or minimized (e.g. lost raw material) to find an optimal control pattern. To calculate extreme values an algorithm based on the Nelder-Mead simplex search [9] is used. This method is characterized to be robust concerning local minima as it doesn't use gradients.

3 Concept for online monitoring of production processes

The key idea of the online monitoring is that the calculated optimal features (which are calculated online during the process) can give a prediction of the process performance. Hence if the predicted process performance tends to be out of the specifications the operators have the opportunity to modify the control structure or process parameters. For an overview about first research topics about online monitoring for production process using Data Mining methods see [6].

The described procedure in section 2 delivers a SVR model based on the relevant features of the process. Furthermore this model is used to calculate the optimal feature values. For online monitoring purposes this model has to be further adapted. A first adaptation is carried out by selecting only interpretable features and features that can be calculated during a production. Furthermore selecting different time segments in which these features are calculated has to follow some requirements as well. Figure 4 gives an overview of the different steps that have to be carried out to make the Data Mining process capable for online monitoring.

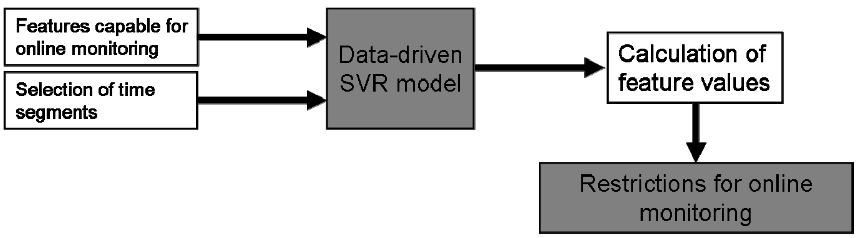

Fig. 4: Calculation of restrictions for online monitoring production processes.

3.1 Feature requirements

Features have to fulfill two major requirements. First these features have to be interpretable already during the production process. For example regression coefficients and features coming from second derivatives of the production data will be neglected. This implies already that a trade-off has to be done between the information included in a feature and its interpretability, as usually more complex features include more information. Furthermore some features can only be calculated if the production or the selected segment is already finished (i.e. mean values). That's why these features can't be used as well for giving restrictions for online monitoring the process. It has to be mentioned that reducing the available features for ranking and model building usually reduces the accuracy and increases the number of selected features. As an example for the behavior of features over time figure 5 shows the behavior of features calculated from a step response coming from a PT2 with a delay time of 1 and a damping of 0.5. By looking at the feature value behavior this shows already the difficulties for an interpretable feature selection for online monitoring. As an example the plotted values of the range and the mean of the second derivative are not intuitively clear and thus it is not sure on how to react if a production runs out of its optimal control pattern and doesn't follow its optimal values anymore. As a conclusion these features have to be removed from the analysis.

3.2 Selection of time segments

Another important requirement is the proper selection of time segments. The features shown in figure 5 only have limited information about the underlying PT2 system as they take into account the whole production process. As example the value of the range reaches its maximum as the PT2 reaches its maximum as well and doesn't change afterwards. Hence after the PT2 has passed its maximum the feature doesn't deliver new information and the feature value can't be used for describing the production further. On account of this the selection of the right segmentation is important for the later use of features for online monitoring. There are two different options for the segmentation of a graph. The first option is to select a fixed time length; the second option is to define

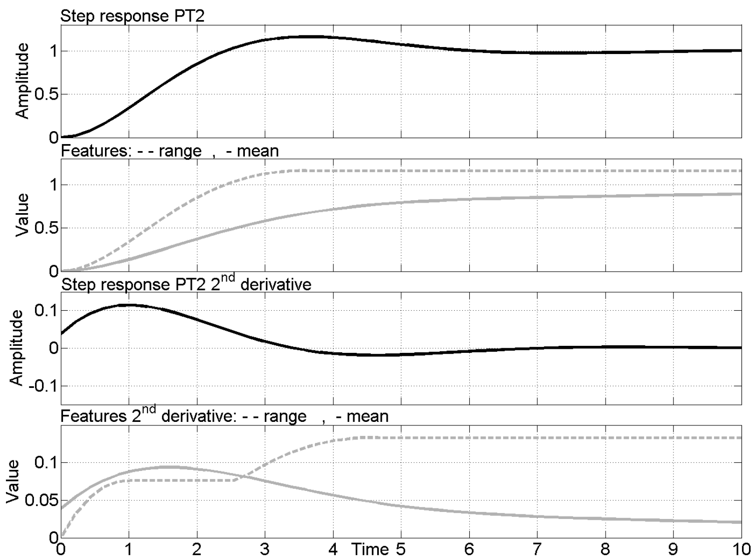

Fig. 5: Examples of features behaviour calculated from process data over time.

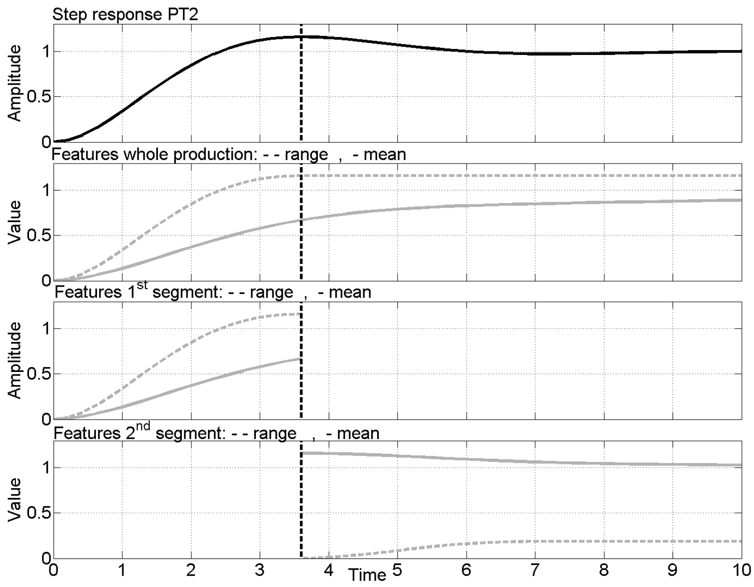

Fig. 6: Features calculated from process data after segmentation

landmarks regarding the production process. Landmarks, like local maxima or minima usually deliver better results as they can take into account more information about the underlying process. Figure 6 shows the prior defined PT2 system where the step response is segmented at the maximum for the generation of new feature values. Features coming from the second segment bring in new information compared to features describing the whole production. From these features the most significant ones have to be selected and their optimal features values can be calculated using techniques as described in section 2.

3.3 Online analysis of feature values

To perform an online monitoring of a production the selected and afterwards calculated optimal feature values can be used. As each feature symbolizes a segment during one production it can give clear restrictions for a production during a segment. Each time one segment is finished, the selected features can be used for the next segment to give restrictions for optimal control. For monitoring a process the generation of features, the selection of a small subset and the calculation of their values helps to focus on the most important process parameters for keeping the production inside its optimal control. The generation of these restrictions and how they can be used for online monitoring the process is shown in the following section on a rheological glass forming process.

4 Application to an industrial glass forming batch process

4.1 Underlying process and data set

The regarded industrial batch process is used for the production of glass tubes out of thick walled blank cylinders [1], [2]. The tubes must meet the desired specifications very precisely. For the production the cylinders are fed into the furnace and thin tubes are pulled out with the velocity v. The furnace is heated up to the temperature ϑ_{oven},. Using the heating power P_{oven}. To achieve the wished diameters D the productions are controlled by ϑ_{oven} and v. The temperature of the glass is measured with the parameter ϑ_{glass} (figure 8).

The investigated process has strong radiation effects and the material properties are changing with time which leads to a nonlinear and time variant behavior. As glass tubes not fulfilling these specifications can't be used the aim has to be to loose as less material as possible during the start-up of a production. That's why as performance index the cumulated lost glass volume is considered until the tube diameter enters a predefined tolerance band close to its nominal diameter *D*. The investigated dataset consists in 850 productions. Before feature extraction the measured time series P_{oven}, ϑ_{oven}, ϑ_{glass} and v, are presampled, low-pass filtered and outliers are removed. Figure 7 shows an exemplary production of one glass tube production. For the analysis two landmarks are defined which leads to 4 different segments plus the whole production process. The following section describes how to generate at first a preliminary model using as much information as available to perform a first analysis of correlations. In a further step a model only based on features that are interpretable and can be used for online monitoring are selected. Finally the optimal feature values are calculated using the prior described simplex algorithm and restrictions are generated for online monitoring the process.

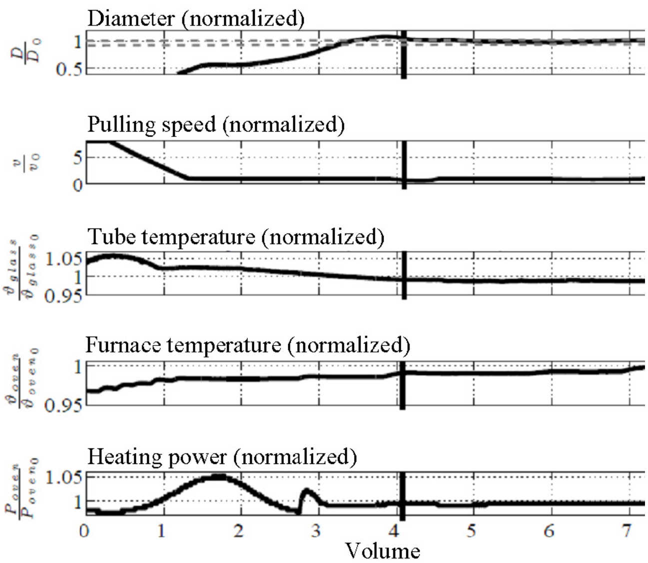

Fig. 8: Exemplary glass tube production. The vertical separation line is placed where the diameter *D* of the glass tube enters the defined tolerance band

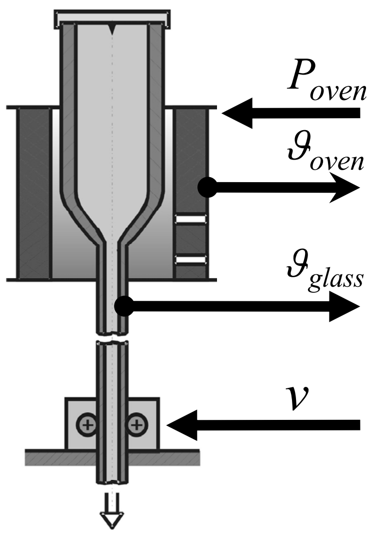

Fig. 7 Industrial batch process for the production of glass tubes.

4.2 Calculation of a preliminary SVR model

First a Data Mining analysis is performed using as much information as possible. This step is necessary to survey if there are any correlations at all between the defined performance index and the measured process parameters.

As this analysis won't be used for online monitoring, artificial features can be used as well and therefore from every time series the first and second derivative is taken. In detail from each process data measured the following six features are extracted: Minimum, maximum, mean value, range and regression coefficients (gradient and offset) coming from a linear regression model.

The results of the fund feature subset as well as the information included in each SVR model are shown in table 1. Using cross-validation techniques the final information included in the SVR model shows a fit of 79 % compared to real process data. This indicates strong correlations between the process data, in depth the tube temperature ϑ_{glass} and the pulling speed v, regarding the defined performance index. The pulling speed is used to control the diameter of the glass tube and that's why the correlations can be explained between the achieved quality and the pulling speed. Interesting is that all other features only depend on the first derivative of the tube temperature ϑ_{glass}. That means that not the actual tube temperature has an influence on the quality of the produced glass tubes but the temperature fluctuation. Regrettably almost only complex features that can't be used for online monitoring a production are selected for model building. Furthermore some features (i.e. features ranked 1 and 2) can only be calculated performing an offline analysis of the productions. Therefore in a next step a SVR model including only features capable for surveying the process online is calculated.

Table 1
Selected feature subset for preliminary SVR model

Rank	**Name of feature**	**Declaring real process (cum)**
1	1st Derivative Gradient of ϑ_{glass} after entering tolerance band	48%
2	1st Derivative Mean v after entering tolerance band	59%
3	1st Derivative Mean ϑ_{glass}	70%
4	Range ϑ_{glass}	75%
5	1st Derivative Maximum v	79%

4.3 Calculation of final SVR model and calculation of restrictions for optimal control

As described in section 3 for the generation of the final model only features and process parameters that can give intuitive results when the process is still running will be used. Regarding the preliminary SVR model only the process parameters v, ϑ_{glass} and the first derivative of ϑ_{glass} are considered for the generation of the SVR model. Furthermore, only minima, maxima and motion ranges will be extracted and the features coming from the segment after entering the tolerance band will be removed as well. Performing a second analysis leads to a bigger SVR model as now these features are less complex and therefore less information is included in each extracted feature. Comparing this model with the prior calculated one it can be seen as well that the new model doesn't reach the same amount of model fit as the prior one although more features are selected. Using cross-validation techniques a model fit results in 69 % compared to real process data. The resulting subset of features is shown in table 2. These features can now be used to calculate their optimal values and finally giving restrictions for optimal process control.

Table 2
Selected features for final SVR model

Rank	Name of feature	Restriction for optimal production (normalized)	Declaring process (cum.)
1	Range ϑ_{glass} before entering tolerance band	< |0.085|	31 %
2	1st derivative Range ϑ_{glass} before entering tolerance band	< |10.1|	39 %
3	Range ϑ_{glass} after Diameter passes maximum	< |0.023|	46 %
4	1st derivative Minimum ϑ_{glass} before entering tolerance band	> -7.3	54 %
5	Maximum v after diameter maximum	< 1.02	58 %
6	Maximum 1st derivative ϑ_{glass} before entering tolerance band	< 5.62	62 %
7	Range v before entering tolerance band	< |0.430|	67 %
8	Maximum ϑ_{glass} before entering tolerance band	< 1.07	69 %

4.4 Using SVR model and features for online monitoring production processes

The calculated SVR-model can now be used for the extraction of optimal feature values. For this the prior described Nelder-Mead simplex algorithm is used. As the lost glass volume during one production should be minimized the algorithm is used to find the minimum of the 8 dimensional surface. For a brief overview figure 11 shows the first two ranked features including the defined performance index. Inside the minimum lies the optimal control pattern where the glass volume is minimized and depending on the values of the features, the production is running optimal or not.

The calculated optimal feature values can now be used for online monitoring the productions. For this restrictions are generated regarding the maximum value one feature can achieve and still stay inside the optimal control pattern. It is defined that a production is still running optimal if its value exceeds at most 20% from the optimal value found through the simplex algorithm. Finally restrictions as shown in table 2 can be drawn to describe if the process is running optimal or not. Figure 9 shows a production where the process data lies outside the restrictions. As the range from ϑ_{glass} is higher than the defined optimal value, it can be already said in the beginning of the production that more glass volume will be lost until the diameter of the tube enters the tolerance band. This leads to a slow settlement of the diameter and thus to a bad quality regarding the defined performance index. For a better view only the first three restrictions are drawn inside the plot. In contrary figure 10 shows a production where the optimal feature values don’t leave the defined restrictions. Therefore this production has a fast settlement regarding the entering of the diameter into the tolerance band.

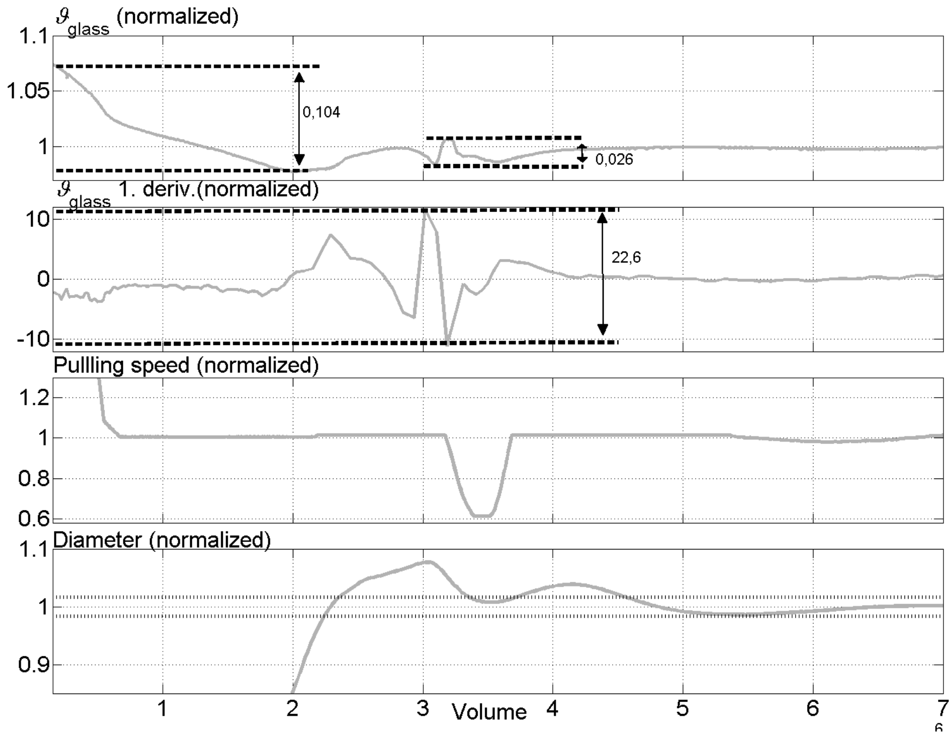

Fig. 9: Production running outside optimal control pattern.

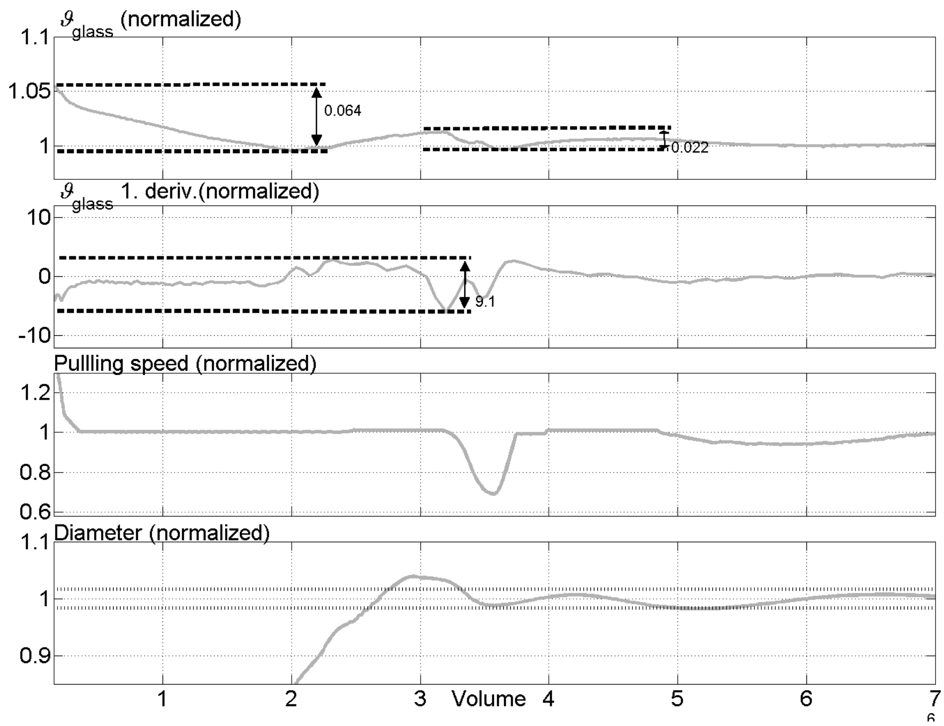

Fig. 10: Production running inside an optimal control pattern.

To confirm the found feature values over a broad range, finally 500 productions were taken from a database and tested if they fulfilled the restrictions during a production or not. Figure 12 shows productions fulfilling the restrictions defined in table 2 and plotted with productions not fulfilling these restrictions. Clearly an improvement is shown. Productions running optimal show a much faster settling towards the defined tolerance band and thus a significant improvement regarding the performance index.

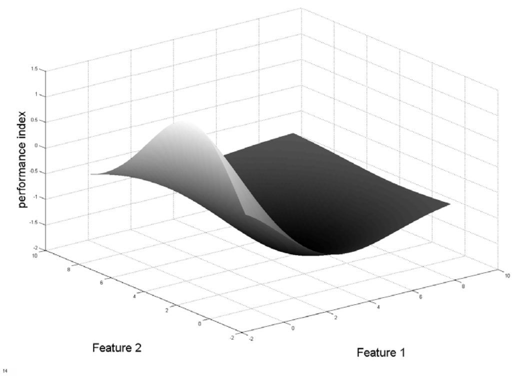

Fig. 11: Impact of the first two ranked features to the performance index

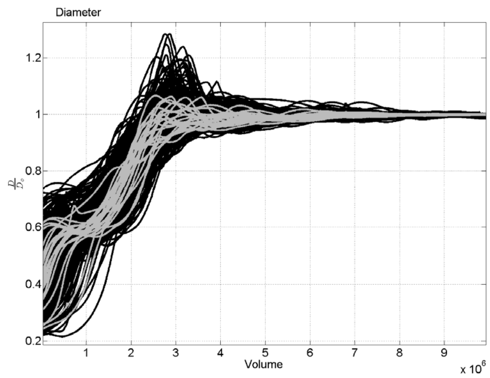

Fig. 12: Comparing productions running outside optimal control patterns (black) with productions fulfilling optimal feature values within a tolerance of 20% (grey)

5 Conclusion

In this paper a new data driven approach of analyzing and online monitoring a complex industrial production process is presented. The idea is in first place to detect hidden and not considered information in an historical database and to make the found knowledge available. Within here the generation of restrictions for online-monitoring production processes is presented. Therefore a prior SVR model is generated including a feature selection and feature ranking procedure. This calculated model including a subset of complex features can then be used to control if there are detectable correlations and interactions between different process parameters. Hereupon a model only based on easy to interpret feature values is calculated. Finally these values can be used in form of restrictions for online-monitoring the process. These restrictions are used to control if the process is running close to an optimal control pattern or not. As performance index the lost glass volume of an industrial glass forming process has been used to find correlated significant subset of features. Finally only a small number of features are needed to give clear advices on how to run the process optimal.

Future work will be concentrated on the steps how the optimal feature values can be used as soft sensors and be used directly for the design of new control concepts. Further research will also be based on a proper analysis of the found SVR model regarding generality (i.e. regarding different produced diameters) and security (i.e. does the model describe the whole feature space).

6 Literature

[1] T. Bernard, E. Moghaddam, *Nonlinear Model Predictive Control of a Glass forming Process based on a Finite Element Model*, Proc. 2006 IEEE International Conference on Control Applications, (2006)

[2] T. Bernard, M. Sajidmann, *A new Fuzzy based multiobjective optimization concept for process control systems*, Proc. 7th Fuzzy Days, International Conference on Computation Intelligence, 2007

[3] C. Kuehnert, T. Bernard, *Extraction of optimal control patterns in industrial batch processes based on Support Vector Machines*, 2008 IEEE Multi-Conference on systems and control, 2009

[4] C. Kuehnert, J. Minx, T. Bernard, H.-B. Kuntze: *Optimierung der Prozessführung komplexer verfahrenstechnischer Prozesse mit Support Vector Machines*, Automatisierungstechniosche Praxis atp 51 (2009), 7, S. 47-49

[5] M. Last, A. Kandel, H. Bunke, *Data Mining in time series databases*, World scientific publishing, 2004

[6] R. Nikhil, J.Lakhmi, *Advanced Techniques in Data Mining and Knowledge Discovery*, Springer, 2005

[7] R. Mikut, *Data Mining in der Medizin und Medizintechnik*, Universitätsverlag Karlsruhe, 2008

[8] R. Mikut, *Data-Mining-Analysen mit der MATLAB Toolbox Gait-CAD*, at-Automatisierungstechnik 56(7), 2008

[9] J. A. Nelder, R. Mead, *A simplex method for function minimization*, Computer Journal, 1965,

[10] R. Otte, V. Otte, V. Kaiser, *Data Mining für die industrielle Praxis*, Hanser Verlag, 2004

[11] A. Rakotomamonjy, *Analysis of SVM regression bounds for variable ranking*, Neurocomputing, Vol. 70, 2007

[12] A. Smola, B. Schölkopf, *A Tutorial on support Vector Regression*, NeuroCOLT2 Technical Report Series, (1998)

[13] J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, *Least Squares Support Vector Machines*, World Scientific, Singapore, 2002

[14] I. Witten, E. Frank, *Data Mining Practical Machine Learning tools and techniques*, Elsevier, (2005)