Development of an Adaptive, Statistical Model Based, Optimal Control Algorithm

Eric Loftsgaarden

Department of Mathematical Sciences
Montana State University

May 3, 2016

A writing project submitted in partial fulfillment
of the requirements for the degree

Master of Science in Statistics

APPROVAL

of a writing project submitted by

Eric Loftsgaarden

This writing project has been read by the writing project advisor and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the Statistics Faculty.

05/05/2016	John Borkowski
Date	John Borkowski Writing Project Advisor
Date	Mark Greenwood Writing Projects Coordinator

Abstract

With the rapid proliferation of sensors in complex manufacturing processes it is becoming increasingly difficult to monitor and efficiently manage all components of the process. A paradigm shift is taking place moving process control out of the hands of expert human operators to automated, data driven control algorithms. Large volumes of high dimensional data present challenges in processing speeds for the automated controllers.

In this paper we present some early exploration of a control process that combines fast low order statistical approximation models with a slower, more robust mathematical model to create an accurate high speed process control algorithm. To develop this algorithm we will use a binary distillation column as the primary reference model. In particular, in this paper we explore the use of localized statistical models and impacts of change point detection on the algorithm.

In order to be able to use a low order statistical model it is important to realize that the model cannot accurately predict over the very large, complex parameter and predictor space. The used of localized low order statistical models is discussed and considerations that are needed in order to keep controller consistency when transitioning between models.

While collecting data for an ongoing process, the underlying statistical model can change due to a perturbation in the system. Examples of this can be found in a wide variety of applications. One such example occurs in a binary distillation column where an inferential control system is used, and a change to the feed composition may impact the effectiveness of the statistical control model meriting a potential change of model. Another example occurs when forecasting demand for a product, and an outside perturbation such as the introduction of a new competitive product potentially affects the current models ability to effectively forecast demand. In both of these cases and in others, rapidly detecting the change point and dynamically selecting a new statistical model can improve the overall process.

In this paper we present a process for dynamically detecting a change point and subsequently selecting a new statistical model for the system from a database of potential models. Considerations for the process are made for speed of identifying a change point and model selection (measured in both computational time and number of data points needed), accuracy of the new model, and potential for type I and type II errors in relation to change point detection.

Contents

1 Introduction

With the rapid proliferation of sensors in complex manufacturing processes it is becoming increasingly difficult to monitor and efficiently manage all components of the process. A paradigm shift is taking place moving process control out of the hands of expert human operators to automated, data driven control algorithms. Large volumes of high dimensional data presents challenges in processing speeds for the automated controllers. Many companies are facing the challenge of taking the large volumes of data that they are collecting and turning it into usable information.

Sensors are now capable of monitoring the performance of nearly every component of a large process collecting data at very short intervals. There are manufacturing processes that generate millions of pieces of information on an hourly basis. In this paper will will explore an algorithm for using this data to improve the efficiency and performance of the manufacturing process.

2 Process

Data-driven control is the process of designing the parameters of a controller based on input and output data without any consideration of the underlying process models. In data driven control, parameter adjustments are made based on volume data rather than on just a few samples. [1] Data driven monitoring methods have become more popular in recent years, especially when it is not practical to develop model based monitoring due to the complexity of modern day processes. With the widespread use of control systems there is a large amount of real-time process data available. This data availability has generated significant investments in new technologies for monitoring and control systems. [2]

Early detection and correction of faults in a process can greatly reduce product loss. Human operators have a tendency to focus on a few key indicators and are often slow to react. When an intervention is taken it is common to overreact to the sensor reading causing oscillations in the process, resulting in a lengthy time to fine tune the adjustments. It is also common for an operator to be conservative in the adjustment only partially correcting the process with each adjustment again increasing the time where the process is out of control.

By automating the control process, identification of a degradation in the process measured across a large number of monitoring devices can be sped up. Automated adjustments can take place in a single step rather than a series of manual steps to achieve the same fine tuning. Both of advantages of automated processing as well as the reduction of human error contribute to the efficiency gains that can be achieved through an automated control process.

3 Binary Distillation

To develop and test the tools needed for the control algorithm a binary distillation process was selected as the reference model. Binary distillation columns are a commonly used and widely studied industrial system. Distillation is a process of separating different chemical compounds from a mixed liquid (feed product) using differing boiling point properties of the compounds. The more volatile component is boiled out of the mixture and then condensed and is referred to as the overhead product. The less volatile component remains in a liquid form and falls to the bottom of the column and is thus referred to as the bottoms product.

To increase the effectiveness of the distillation process a temperature gradient is created inside the distillation column with many zones where an equilibrium between the product is established. Each of these zones are called trays or distillation plates as in figure 1. Each individual tray can be modeled by a differential equation. The reflux of the overhead product and the input of heat from reboiler at the bottom of the column can also be modeled. Thus creating a complex system of coupled differential equations that can be used to accurately describe the distillation column.

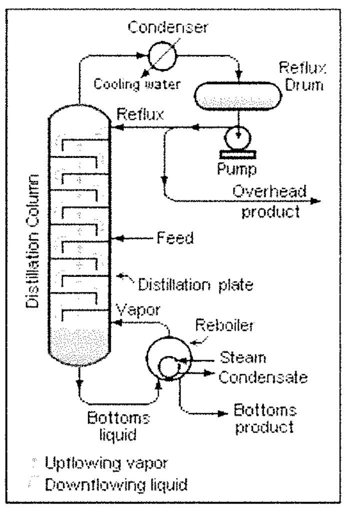

Figure 1: Schematic of a binary distillation column

In binary distillation processes the product composition of the overhead and bottom products are key to the overall effectiveness. Accurate on-line measurements of product compositions are very difficult. Most of the measurements are completed off-line using gas chromatography or near-infrared analyzers. These systems are expensive to purchase and maintain, in addition they introduce significant lag into the measurement process. [3] This time lag can result delays in detecting a drop-off in product quality introducing additional costs into the manufacturing process and reducing overall end product quality.

Kano et al. [3] indicate that tray temperature control is often used as a proxy for product composition. In more dynamic distillation columns where flow rates, feed composition or internal pressures can fluctuate, temperature control fails to be an accurate enough predictor of product composition. Including additional variables in a predictive model is necessary for improving product compositions control.

Often the distillation columns contain a large number of temperature trays often 30-50 or possibly more. The measurements on one tray are highly coupled with measurements from adjacent trays. The response also has a time dependent component as changes in the process do not immediately result in the system obtaining a new steady state. Partial least squares or other dimension reduction approaches are often used in models for solving this challenge. [3]

Binary distillation provides a model that is simple enough that it can be simulated to a high degree of accuracy in a reasonable time frame providing generated "real" data for the testing of the proposed algorithm. It is also a complex enough process that numerous simplified models have been developed as prediction models. It also provides enough of the complexity of larger systems that meaningful time difference for various approximation algorithms can be compared. All of these attributes make the binary distillation process a good fit for the development and testing of the proposed process control algorithm.

In the binary distillation process the purity of both the top product and the bottom product will server as response variable. The actual responses are typically measured with a significantly lower frequency than the predictor variable which can be monitored in near real time. There are two types of predictor variables in the system. The first type are sensor data which are purely input only measurements of various parts of the system. In this case primarily the temperature at each plate and the concentration of the feed product. There are also adjustable predictor variables which the operator can change to affect the process. These are the primary variables that the feedback controller would adjust to keep the process in control. For the distillation column these would include the reflux rate, the reboiler temperature, and the feed rate. Emphasis will be put on the adjustable predictor variables and they will be included in every model in order to provide the automated controller the maximum amount of information and flexibility to keep the process in control.

4 Adaptive Model

Borkowski et. al. [4] have proposed an adaptive data driven computational framework that blends real time sensor data, high-fidelity phenomenological-based models(HFMs) and a library of statistical approximation models (SAMs) to select an appropriate SAM to predict the state of the physical process, in this case the binary distillation column. In this paper we will explore some of the challenges identified in creating the adaptive statistical model.

The suite of HFMs provide highly accurate, high dimensional models of the physical plant. These models can provide very detailed information on the physical process. The level of dimensionality and accuracy in these models comes at a high computational cost resulting in a speed that is much slower than the real time data. In order to compensate for the slow speed of the HFM models, a suite of statistical models will be developed to work in conjunction with the HFM's. The SAMs provide a low dimensional model of the system response but are only locally accurate. The dynamic SAM sub-process highlighted in red in Figure 2 will be developed and evolved to be locally accurate and to anticipate drift in the predictor and parameter space. By leveraging the HFM models to simulate data in the direction of projected process drift locally accurate candidate SAMs can be developed before the physical plant shifts to the new portions of the predictor space space. Once the plant has entered the new area the best candidate SAM will be selected while continuously monitoring other candidate SAMS for a reduced error model. In this way the statistical model can "stay ahead" of the physical process.

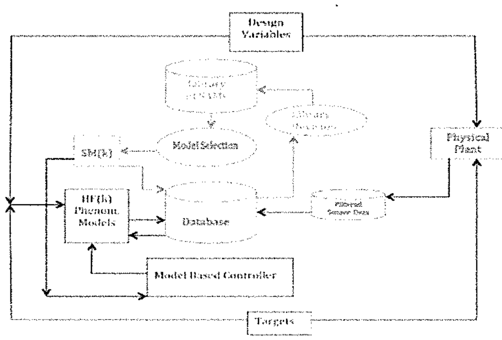

Figure 2: Proposed computational framework

The current "best" SAM will be linked with the feedback controller to maintain process stability. Additional work, outside the scope of this paper, will be completed to explore different controller configuration in order to ensure that closed loop stability is maintained. As the process drifts outdated models will be archived for future use when the process re-enters the neighborhood where each model is locally accurate.

Database

New data will be continuously generated from both the HFMs and from the actual process monitoring devices. The process will have a continuous flow of censor data, response data and quality control data. This data will be added to the database to be used as input data for generating the localized candidate statistical models.

The validity of older data will be process and potentially data stream dependent. Each macro level process and each data stream will be evaluated separately to determine how to handle older data. A few of the potential treatments for handling historical data include:

- Removal after a certain age: At some point data may become unreliable for predicting the current process state.
- Down-weighting: Current data may be more influential than historical performance so some sort of down weighting of historical data may be appropriate. One example may be using an exponentially weighted model.

- Maintained indefinitely: This data my be reliable for as long as the system operates. Therefore, we want to keep and use the data for the foreseeable future.

- Combination: Data may fall into more than 1 of these categories. For example, once a data item hits 3 weeks old we may start down-weighting it and then once it hits 3 months of age we may then remove it as unreliable.

Library Revision

The library revision process will have three primary components. First, new data will be used to adjust the existing model parameter estimates in real-time. Second, the existing localized model library will be evaluated along with any newly generated candidate models. If a new candidate model is performing better than an existing model in the model library it will be promoted to replace the poor performing model. Third, the direction of predictor drift will be anticipated and new candidate models will be generated for the nearby unexplored portion of the parameter space. A particular focus will be made to ensure that the adjustable predictors are sufficiently explored in order for the controller to be able to effectively keep the process within specification.

Generating new models will require some decisions about the types of models that we want to consider in the process. Higher order models will be able to cover a broader portion of the parameter space but require more data and time to provide an accurate model. First order or low order models are easier to select but are more sensitive to drift. Automated stepwise variable selection can be used to identify candidate models for evaluation.

Having the library revision process in place will also enable the system to easily adapt to changes in the physical plant. For example the addition of a new sensor will be accommodated by adding the data to the database and then making the new parameter available in the model selection process. If this new sensor proves valuable it will be included in new candidate models and promoted to the model library. Likewise if a sensor become unreliable because of wear and tear, noise or other reasons it will naturally and automatically be removed from the SAMs. In this way the algorithm will easily and often automatically adjust to system changes without the need to incur costly shutdown and restart processing.

Library of SAMs

The library of SAMs will be a pre-determined subset of the best performing candidate models in a particular parameter and predictor subspace. Every time a new measurement of the response variable is received a statistical quality control check (SQC) will be performed on all models in the model library using a standard SQC process monitoring procedure. Data from the phenomenological models will also be used to perform quality checks on the models in the library.

If some models in the library are identified as poor performing models in the local parameter space then they will be removed from the library and replaced with better candidate model from the library revision process.

Model Selection

From the library of models a consensus "best model" will be selected and promoted as the active model for the controller. This model will be selected based on one or more common models selection techniques. For the studies in this paper the Akaike Information Criteria(AIC) was used as the selection criteria. Future studies may included a broader spectrum of criteria.

Future consideration will also be given to using a linear combination of models, especially as we near the boundary of the predictor space of the given localized model. This will help facilitate the transition between models as the process drifts in the predictor space.

SM(k)

The "SM(k)" step indicates the best statistical model at time step k of the process. This is the best model selected for the localized parameter step at a given time. This model will be promoted to the model based process controller. This consensus model can then be used to recommend adjustments to the manipulated variables in conjunction with existing targets.

5 Local Approximation

Finding a statistical approximation model for a system that has a very large high dimensional parameter space in a system with sparse data and stringent time restrictions is challenging. Zhang et. al. [5] have described a process for reducing the dimensional complexity of a problem by using a local linear approximation method. This method provides a precise approximation with low mean squared errors while also reducing the time complexity of the problem compared to the traditional differential equation solver approach. Efficiency improvements of 20 fold were found using this method. Approaches such as this are a very active area of research as the need increases for handling very large data volumes efficiently. Further research in this area will be needed to fully implement the proposed model using localized low order approximations.

To illustrate how this process will work we examine a single predictor scenario where the true underlying model is of the form e^{x^2}. Data from the model $y = e^{x^2} + error$ was simulated by generating random x values from a uniform distribution and $errors$ from a normal(0,.03) distribution. Figure 3 was created by fitting a linear model to the generated data when the predictor variable was in the interval [.2,1]. In this range a linear fit provides a good approximation of the resulting data. The best fit line has been plotted on this diagram. As the predictor space expands to the interval [0,2] we can see in Figure 4 that this original linear model, shown in red, is no longer a good predictor model for the responses. Even the updated linear model no longer provides a very good model for the data. In this scenario we would need to fit a cubic polynomial in order to provide a good polynomial approximation of the real model. In a large parameter space increasing the order of the polynomial model significantly increases number of terms and thus the amount of data required to fit the approximation. Even a cubic polynomial in a high dimensional space becomes unusable for the practical problem of fitting the underlying exponential model being considered.

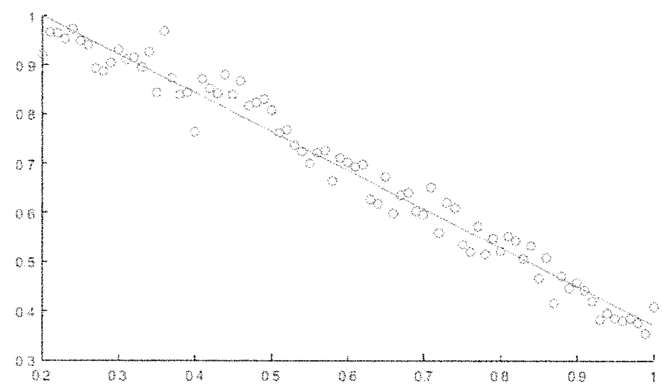

Figure 3: localized linear approximation

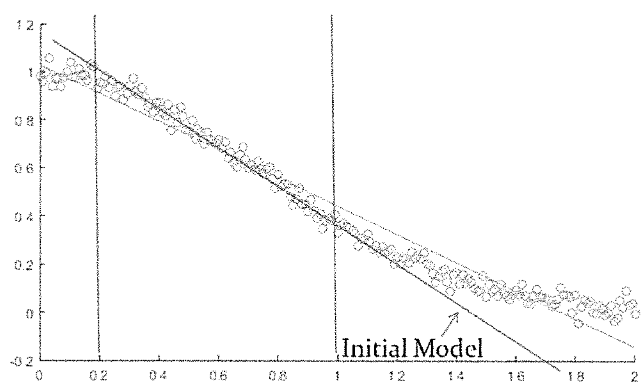

Figure 4: localized linear approximation of the same true model over a slightly larger region of the predictor space

Figure 5 shows another scenario where our parameter space has drifted to the interval [-1,1]. In this scenario a quadratic polynomial provided a good fit to the actual data. Rather than continuing to fit increasingly higher order polynomial models as the predictor space expands, we will fit low order models to localized areas of the predictor space in order minimize the amount of data needed and to keep the computational speed high. Figure 6 provides an example of this where the outside edges have fit linear models and the peak of the curve has been fit with a quadratic model. As we drift to different portions of the parameter space the best localized model will promoted to be the primary model for the controller. As the process enters the model overlap area shown in Figure 6, both models will be active in the model library and the best model will be selected based on a minimum mean squared error. Having both models available in the model library will allow the system to easily and efficiently swap to the new model reducing and potentially eliminating the risk of the controller making an incorrect adjustment.

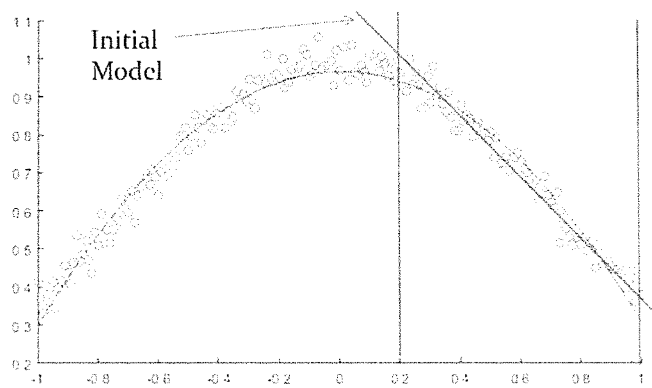

Figure 5: localized linear approximation of the same true model over a slightly larger region of the predictor space

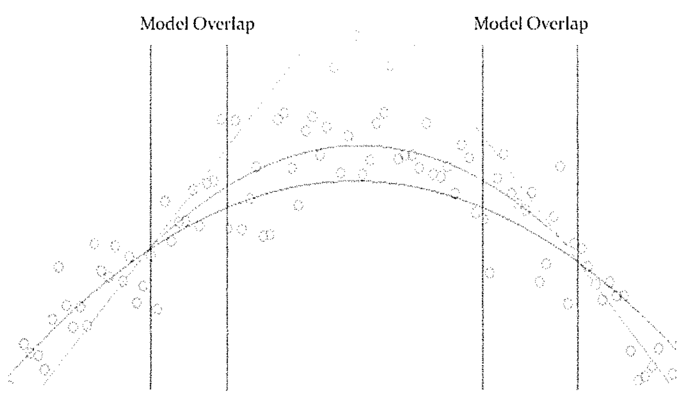

Figure 6: Localized low order approximation model with overlap between neighboring segments

6 Change Point Detection

Given a pre-defined acceptable error level in the prediction model we can turn to traditional control charting techniques such at the Shewhart, CUSUM or EWMA procedures to identify when our localized model is no longer performing adequately. We can also rely on extensions of these procedures to the multivariate predictor scenario. These multivariate methods are based on statistical methods such as Principal Component Analysis (PCA) and Partial Least Squares (PLS). MacGregor et. al. [6] provide an overview of these techniques and their use in industrial processes.

An out of control sign in the process can be caused by a number of items. It could be that we have drifted out of the local predictor space where our current model is valid in which case we would want to select a new model for the new part of the predictor space that we are in. Second, it could be that the model is inadequate and needs to be adjusted. To address this the other candidate models will be evaluated or potentially new candidate models would be generated from the library revision process. The third scenario that could occur would be triggered by an underlying change in the "true" model. Generally this will happen if there is some kind of physical change in the system. One example of a change would be a sensor that is starting to fail and is producing unreliable readings. If this sensor was included as a predictor in the model, the model would no longer be valid and other candidate models (ones without the faulty sensor) would perform better and thus be promoted. Lastly we could have a scenario that we have a false out of control signal. Depending on the parameters used in the control charting process we would expect this to happen at an average, predetermined frequency. In this case the current model may still be the best model. Keeping model changes to a minimum if the underlying local model has not changed will be important for the stability of the system. Further research is needed to identify an automated method to quickly determine if our our out of control signal is a false alarm so we can leave the current model in place without unnecessary transitions.

For the proposed algorithm an out of control signal will be the catalyst for a number of processes. It will immediately trigger a model selection and the potential promotion of another model in the model library. As discussed above, one of the primary goals of the HFM models is to anticipate the direction of the predictor drift and have candidate models ready for promotion. In the case where there are no valid candidate models the process will move from a stable state and enter a transitional state. Figure 7 shows the transitional state. When the process is in the transitional state the controller will be blind, i.e. there is no adequate model for predicting the response variable. One of the primary goals of the algorithm and the use of the HFM models is to minimize the time in the transition state. Further research will be conducted on the creation of an objective function to tune the effectiveness of the algorithm and time spent in the transition state will be a large contributor to the objective function.

When the process is in the transition state a continuous search for new models will take place with the arrival of each new data point and weighting the data so that the newest data has the greatest influence on the model. A simulation study has shown that initially the "Best" model will change after nearly every new data point, but as more post change point data is collected, the rate of model change will slow down rapidly until a new stable state is reached. Another area for further investigation would be to better quantify the number of incoming data points needed

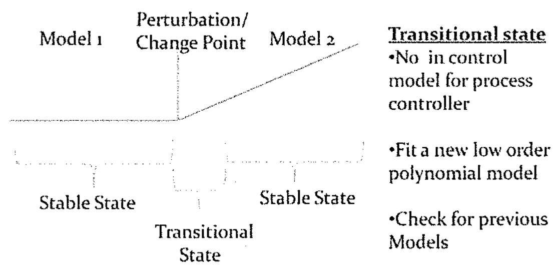

Figure 7: Transitional State

to reach the stable state as a function of the number of potential predictors, the number of actual predictors used in the model, and the signal to noise ratio.

7 Items for Further Study

Initial background research has raised a large number of questions that need further exploration in order to fully understand, develop and quantify the effectiveness of the proposed process. At this point we haven't encountered any barriers that appear to be unsolvable. A few of the key items that need additional research are:

- Development of an objective function to be used to compare algorithms
- Multivariate control tracking
- Resetting and re-balancing of control charting
- What is the best model selection criterion to use? AIC? something else?
- How do we develop and decide which model generation methods to use? Linear, quadratic, poly, interactions, non-linear etc.?
- Do we use naive model generation methods (purely data driven) or try to develop smart methods (data combined with process knowledge)?
- How do we forecast what area of the parameter space we want to explore next?
- How do we use experimental design techniques to direct the HFM to the optimal sampling plan?

Bibliography

Sanfelice Bazanella, Alexandre, Lucíola Campestrini, and Diego Eckhard (2012). *Data-Driven Controller Design.* ISBN: 978-94-007-2299-6. DOI: 10.1007/978-94-007-2300-9. URL: http://cds.cern.ch/record/1501891$%5Cbackslash$nhttp://www.springer.com/engineering/robotics/book/978-94-007-2299-6$%5Cbackslash$nhttp://link.springer.com/10.1007/978-94-007-2300-9.

Zhiqiang, Ge and Song Zhihuan (2013). *Multivariate Statistical Process Control Process Monitoring Methods and Applications*, pp. 1–193. ISBN: 9781447145127.

Kano, Manabu et al. (2000). "Inferential control system of distillation compositions using dynamic partial least squares regression." In: 10, pp. 157–166.

Borkowski, John, Lisa Davis, and Karlene Hoo (2015). "Validating an Evolutionary Framework that Blends High-Fidelity with Data Driven Models and Real-time Sensor Data to Control Dynamic Behavior of Big-Scale Systems." In: *EAGER GRANT PROPOSAL*, pp. 1–8.

Zhang, Peng, Esra Ataer-Cansizoglu, and Deniz Erdogmus (2012). "Local linear approximation of principal curve projections." In: *IEEE International Workshop on Machine Learning for Signal Processing, MLSP.* ISSN: 21610363. DOI: 10.1109/MLSP.2012.6349764.

MacGregor, J. F. and T. Kourti (1995). "Statistical process control of multivariate processes." In: *Control Engineering Practice* 3.3, pp. 403–414. ISSN: 09670661. DOI: 10.1016/0967-0661(95)00014-L.