# Prediction of Batch-End Quality for an Industrial Polymerization Process

Geert Gins[1], Bert Pluymers[2], Ilse Y. Smets[1], Jairo Espinosa[3], and Jan F.M. Van Impe[1]

[1] BioTeC, Department of Chemical Engineering, Katholieke Universiteit Leuven, W. de Croylaan 46 PB 2423, B-3001 Heverlee (Leuven), Belgium
{geert.gins,ilse.smets,jan.vanimpe}@cit.kuleuven.be
[2] IPCOS NV, Technologielaan 11-0101, B-3001 Leuven, Belgium
bert.pluymers@ipcos.be
[3] Facultad de Minas, Sede Medellín, Universidad Nacional de Colombia, Cra 80 No. 65-223 Bloque M8 of 112, Medellín, Colombia

**Abstract.** In this paper, an inferential sensor for the final viscosity of an *industrial* batch polymerization reaction is developed using multivariate statistical methods. This inferential sensor tackles one of the main problems of chemical batch processes: the lack of reliable online quality estimates.

In a data preprocessing step, all batches are brought to equal lengths and significant batch events are aligned via *dynamic time warping*. Next, the optimal input measurements and optimal model order of the inferential *multiway partial least squares* (MPLS) model are selected. Finally, a full batch model is trained and successfully validated. Additionally, intermediate models capable of predicting the final product quality after only 50% or 75% batch progress are developed. All models provide accurate estimates of the final polymer viscosity.

**Keywords:** Industrial batch process, quality prediction, Partial Least Squares.

## 1 Introduction

In chemical industry, batch processes are widely used for flexible production of high-value products (e.g., specialty polymers, pharmaceuticals, and biochemicals). Batch processes are characterized by a fixed recipe, which prescribes a set of processing operations over time. The recipe is followed as closely as possible to ensure a satisfactory product quality. Monitoring batch processes is difficult due to the lack of available online product quality measurements. In most cases, quality measurements are only available after batch completion, often hours late. This makes quality control very difficult: only minor –if any– corrective actions can be taken after batch completion. In the worst case, the just-produced (off-spec) batch must be wasted and the batch run again.

Hence, there is definitely a need for inferential sensors, capable of predicting the final product quality during the batch run, enabling a close monitoring of

P. Perner (Ed.): ICDM 2011, LNAI 6870, pp. 314–328, 2011.

the production process, through which off-spec batches can be detected in an early stage. This early prediction allows corrective actions to be performed during the batch if the expected final product quality is not within specifications. Consequently, less off-spec batches are produced, saving valuable production time, lowering operational costs and reducing waste material. Multivariate statistical methods, originally designed for monitoring continuous processes, have successfully been extended to batch processes [5,6,14,15,16,17,18,20], but actual industrial validation is rare.

**In this work, an inferential sensor for predicting the batch-end quality of an *industrial* polymerization reaction is constructed and validated using such multivariate methods.** The industrial installation is described in Section 2. Section 3 details the data pre-processing using a hybrid *derivative dynamic time warping* scheme, and Section 4 discusses the identification of a *partial least squares* model on the industrial data. To allow online estimations of the final product quality, intermediate models are trained in Section 5. The inferential sensors are validated in Section 6. Final conclusions are drawn in Section 7.

## 2 Industrial Installation

The industrial batch reactor studied in this work is depicted in Figure 1, and a schematic overview of the different events during the production process is given in Figure 2. For reasons of confidentiality, specific process details are not disclosed, and all results will be made dimensionless.

Before the batch run, the raw materials are stored and mixed in the premix vessel $P$. Approximately half of these raw materials are then fed to the batch reactor $R$ in the first loading phase. Next, hot water is circulated through the spiral $S$, heating the reactor content. When the reactor reaches a specific temperature, the initiator activates and the polymerization reaction starts. Shortly after the start of the reaction, the remaining premixer content is added to the reactor in the second loading phase. During the course of the batch, gasses rising

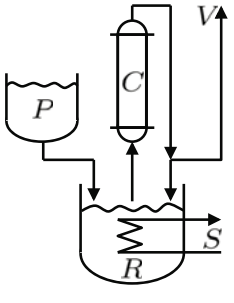


**Fig. 1.** Schematic view of the industrial batch reactor, consisting of a premixer $P$, reactor $R$, condenser $C$, cooling or heating spiral $S$ and gas vent $V$

up from the reacting mixture are condensed in the top cooler $C$, and recirculated to the reactor. Uncondensed gasses escape through the vent $V$, but this process stream is negligible. To keep the reaction rate at an acceptable level, new initiator is added by the operator after approximately 25%, 50% and 75% of the total production time $t_f$. Finally, the batch is terminated and cooled down, and the reactor is emptied.

Measurements are started as soon as the premixer is loaded. Sometimes the premixer is prepared as soon as it becomes available, and sits idle for several hours, while another chemical reaction is still running in the reactor. Other times the premixer is prepared just prior to the start of the first loading phase. Hence, the time between the loading of the premixer and the first loading of the reactor (i.e., the actual start of the process) varies greatly from batch to batch. After the batch is terminated, logging is stopped manually by the operator during the cooldown of the reactor, before the polymer is removed from the reactor.

As a result, the amount of available data varies greatly from batch to batch. This variation is mainly caused by the difference in (*i*) the duration between the loading of the premixer and the start of the first loading phase, (*ii*) the duration of the heating phase, and (*iii*) the amount of time the measurements continue after the start of the cooldown phase. If these phases are not taken into account, all batches have comparable durations. Therefore, the data set is constructed using the detection of the polymerization reaction as the first point. The final point of the data set is the moment at which the cooldown starts. During this period, 30 sensors record various temperatures, pressures, flow rates, and weights; approximately 2000 samples are available for every sensor.

The retained part of the batch operation can be divided in six stages based on the batch recipe. The first stage runs from the detection of the polymerization until the start of the second loading step. The second stage coincides with this second feeding step, after which the third stage starts. The initiator shots are the transitions to the fourth, fifth and sixth stage. From the measurement data, however, not all stages can be identified. At the end of the second feeding step, a temperature drop is sometimes, but not always, observed in the vapor temperature. Both events mark the end of the feeding step but do not always coincide. This makes the unambiguous determination of the end of this stage impossible. Therefore, the stage transition is not taken into account. Furthermore, the first initiator shot (occurring at $t_f/4$) is not observed in all batches

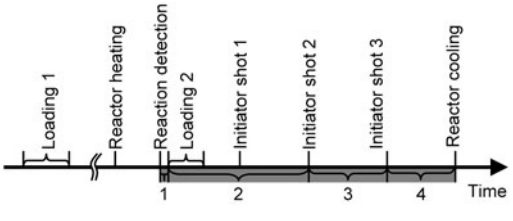


**Fig. 2.** Schematic overview of the batch recipe and the four identified phases

and is likewise discarded. The final result is that the following four *phases* are retained: (*i*) from the detection of the polymerization reaction until the start of the second feeding phase, (*ii*) from the start of the second feeding phase until the second initiator shot, 4 hours after the reaction detection, (*iii*) between the second and third initiator shots, and (*iv*) from the third initiator shot until the start of the reactor cooling. The first phase lasts only a few minutes, while the second batch phase lasts approximately $t_f/2$. Each of the following two batch phases take about $t_f/4$ to complete. A schematic overview of the batch recipe events and the final four identified batch phases is given in Figure 2.

Data of 72 batches is available for training. For each batch, the polymer's viscosity is measured offline via lab analysis. The viscosity is upper and lower bound by specification. Of the 72 training batches, two exhibit a too high viscosity. The validation set consists of 10 additional batches.

## 3 Data Preprocessing

Before a mathematical model can be identified on the measurement data, all batches or profiles are required to be of identical length. Furthermore, similar events should occur at the same moment in all batches in order to improve model performance. First, the data alignment technique is explained in Section 3.1, after which the alignment results are discussed in Section 3.2.

### 3.1 Data Alignment Procedure

*Dynamic time warping* (DTW) has been adopted with success in various fields of research to align measurement profiles of different lengths [1,3,6,11,12,21]. The variant *derivative DTW* (DDTW) has been demonstrated to yield fewer warping singularities [12].

In (D)DTW, the difference between two profiles (i.e., a test and a reference profile) is minimized by dynamically stretching and/or compressing the time of the test profile. This nonlinear transformation is obtained by first constructing the distance matrix $\mathbf{D}$ between the test profile $(x_1\ x_2\ \ldots\ x_M)$ and reference profile $(y_1\ y_2\ \ldots\ y_N)$. The original DTW algorithm uses the Euclidean distance measure, while DDTW uses the difference between the derivatives of both profiles as distance measure [12].

$$\mathbf{D}\,(m,n) = \left( \left.\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\right|_m - \left.\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}t}\right|_n \right)^2 \tag{1}$$

Next, the warping path $\mathcal{P}$ is defined as the continuous path of $P$ different $(m,n)$-pairs from $D(1,1)$ to $D(M,N)$ which minimizes the total distance between both profiles. The warping path is often subject to local slope constraints, a Sakoe-Chiba adjustment window [23], or an Itakura parallelogram [10].

$$\mathcal{P} = \arg\min_{\mathcal{P}} \left\{ \sum_{p=1}^{P} \mathbf{D}\,(m_p, n_p) \right\} \tag{2}$$

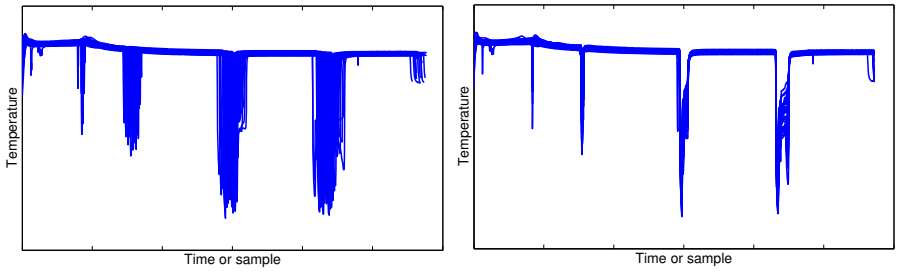


**Fig. 3.** (a) Original profiles of the reference temperature measurement. Significant batch events clearly occur at different times, and all batches have different lengths. (b) Aligned reference temperature measurement profiles. All temperature jumps clearly occur at the same moment, and all batches have an equal length. The axes are unlabelled for confidentiality reasons.

Taking the numerical derivative of noisy data is inherently unstable. To solve this issue, a *hybrid DDTW* (HDDTW) scheme is proposed: a piece-wise linear approximation of the measurement profile is used to compute its derivative. Because the derivative now has piecewise constant shape, the distance matrix $D$ contains rectangular zones where the distance between test and reference profile remains constant. When the warping path passes through such a zone, it follows a diagonal path, resulting in a local linear resampling of the test profile.

The main features of a profile are characterized by rapid changes in derivative, while featureless zones are characterized by an approximately constant derivative. The discretization intervals of the piecewise approximation are very short in feature-rich zones, and the very high warping resolution of traditional (D)DTW is obtained. In featureless zones (e.g., a long period with a constant level or a gradual increase), the discretization intervals are much longer. The net result is a lower warping resolution in these zones and a much simpler (linear) warping. This makes the procedure more robust with respect to measurement noise.

### 3.2 Data Alignment Results

Before HDDTW is applied, a reference variable and reference trajectory are selected for each phase. All events are clearly observable in the vapor temperature at the entrance of the top cooler. Hence, this variable is selected as the reference variable. A representative batch is taken as the reference profile.

Next, each batch phase is aligned using the hybrid DDTW algorithm described in Section 3.1. The warping paths for each of the four batch phases are combined into a single global warping path, which is used to align all measurement profiles for each batch. The parameters for the Sakoe-Chiba, Itakura and local slope constraints for the DTW algorithm are determined using process knowledge. It is observed that the size of the Sakoe-Chiba adjustment window can be set significantly smaller than the commonly used value of 10% of the profile length. This observation corroborates the results reported in [22].

Figure 3(a) displays the vapor temperature for all batches. As can be seen, the major events (temperature drops) occur at different times. Figure 3(b) depicts the aligned vapor temperatures at the top cooler entrance for all batches. Clearly, all batches now have identical lengths, and all major events coincide.

Finally, the data alignment procedure described in Section 3.1 is assessed. The warping path for the third phase of one batch is depicted in Figure 4; similar warping paths are obtained for the other batches and batch phases. In the beginning of the phase, where the temperature measurement profile exhibits a clear drop (see Figure 3(b)), the warping path indicates a more complex warping. This is required to match the new temperature profile as closely as possible to the reference. A simple linear resampling is obtained in the relatively featureless zone near the end of the phase, as evidenced by the linear relation between warped and original time.

## 4 Model Identification

The structure of the mathematical model used for the inferential sensor is discussed in Section 4.1. The selection of the optimal inputs and optimal model order for this mathematical sensor is detailed in Sections 4.2 and 4.3. Finally, the inferential sensor is trained in Section 4.4.

### 4.1 Partial Least Squares Modelling

A *multiway partial least squares* (MPLS) model is used to infer the relationship between the online process measurements and the quality measurements [13,18,20,24]. MPLS is an extension of basic PLS [8]. It is able to handle the three-dimensional data matrices (tensors) characteristic for batch processes. As shown in Figure 5, the data tensor $\underline{\mathbf{X}}$, consisting of $I$ batches with $J$ sensors per batch and $K$ samples per sensor, is unfolded to a $I \times JK$ data matrix $\mathbf{X}$ [18,19].

Before the model is trained, both $\mathbf{X}$ and $\mathbf{Y}$ are mean centered and normalized to unit variance. This centering around the nominal trajectories removes the major nonlinear behavior of the process from the data [18,19].

PLS is a latent variable modelling approach, which decomposes the matrices $\mathbf{X}$ and $\mathbf{Y}$ into $R$ latent variables that each describe an aspect of the batch operation relevant to the final product quality.

$$\begin{cases} \mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E}_\mathbf{X} \\ \mathbf{Y} = \mathbf{T}\mathbf{Q}^T + \mathbf{E}_\mathbf{Y} \end{cases} . \tag{3}$$

The $I \times R$ scores matrix $\mathbf{T}$ is the low-dimensional approximation of the input space $\mathbf{X}$. The $JK \times R$ matrix $\mathbf{P}$ and $L \times R$ matrix $\mathbf{Q}$ are the loading matrices in in- and output space, respectively. The matrices $\mathbf{E}$ represent the residuals.

The projection of the input space $\mathbf{X}$ onto the scores space $T$ is obtained as

$$\mathbf{T} = \mathbf{X}\mathbf{W}\left(\mathbf{P}^T\mathbf{W}\right)^{-1} . \tag{4}$$

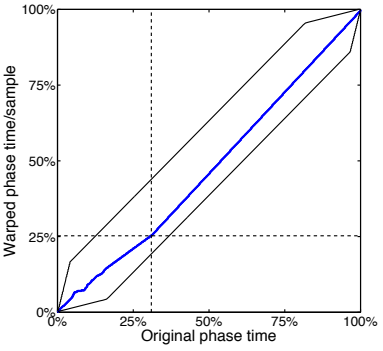


**Fig. 4.** Warped time profile for the third phase of one batch. The solid lines (—) indicate the Sakoe-Chiba and Itakura constraints; the transition between eventful and featureless zones is marked by the dashed lines (- -). Other batches exhibit similar profiles.

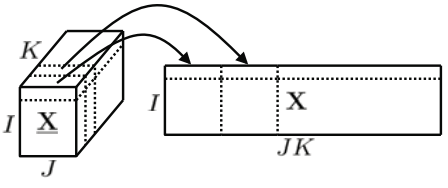


**Fig. 5.** The original $I \times J \times K$ data tensor $\underline{\mathbf{X}}$ of $I$ batch runs, $J$ sensors and $K$ samples is unfolded into a $I \times JK$ data matrix $\mathbf{X}$

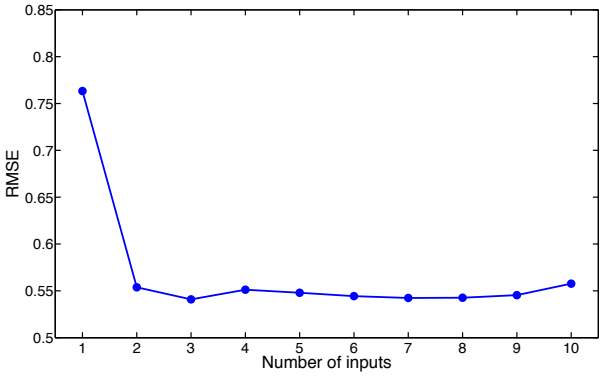


**Fig. 6.** Average crossvalidation error as a function of the number of inputs for the optimal input set of (*i*) the temperature of the reactor, (*ii*) the temperature of the vapor leaving the reactor, and (*iii*) the vapor temperature at the entry of the condenser

The $JK \times R$ weight matrix $\mathbf{W}$ has orthonormal columns such that $\left(\mathbf{P}^T\mathbf{W}\right)^{-1}$ is upper triangular, with ones as diagonal elements.

Via the covariance matrix of the measurement trajectories, MPLS utilizes not only the deviation of each process variable from its mean trajectory, but also the simultaneous and temporal correlation between the measurements [19].

### 4.2 Optimal Input Selection

To select the optimal inputs for the MPLS model, a *forward branch-and-bound* technique is used. First, $J$ MPLS models are trained, each using one possible measurement variable as input and with a number of latent variables $R$ between 1 and 10. Crossvalidation is used to find model with the best overall performance. To reduce the computational requirements, 10-fold crossvalidation is used here. Model performance is characterized by the *root mean squared error* (RMSE).

$$\text{RMSE} \triangleq \sqrt{\frac{1}{I}\sum_{i=1}^{I}\left(\hat{y}_i - y_i\right)^2} \tag{5}$$

The final product viscosity for batch $i$ is $y_i$, the model prediction is $\hat{y}_i$.

Next, the input variable of the best single-input model is combined with each of the remaining input variable candidates. On each of these $(J-1)$ input variable pairs, a new MPLS model is trained and crossvalidated. Again, the inputs for the model with the best performance is retained. This process is repeated until all input variables are ranked from most to least important. Because the batches are randomly distributed into training and validation subsets, the ranking of the input variables and the optimal number of inputs varies between each run. Therefore, the selection is performed multiple times, and the variables scoring the best over all runs are selected as the optimal model inputs.

Two different sets of 3 input variables are retained via this procedure. Both sets share the reactor temperature and the temperature of the vapor leaving the reactor. The third input variable is either the vapor temperature at the entry of the condenser or the flow rate of the cooling water in the condenser. In order to discriminate between both sets, their crossvalidation performance is compared. The former set has a crossvalidation RMSE of $0.533 \pm 0.015$, while the latter has an RMSE equal to $0.605 \pm 0.015$. Based on these observations, the following three input variables are retained: (*i*) the temperature of the reactor, (*ii*) the temperature of the vapor leaving the reactor, and (*iii*) the vapor temperature at the entry of the condenser. Figure 6 depicts the evolution of the validation error for an increasing number of model inputs.

A physical interpretation can be provided for the selection of these input variables. The reaction rate is directly tied to the reactor temperature, and influences the polymer's viscosity. The difference between the two vapor temperatures is an expression of the amount of heat removed from the batch reactor. The total amount of heat removed from the reactor is, in steady state, equal to the amount of heat produced by the polymerization reaction. Therefore, this is an indirect

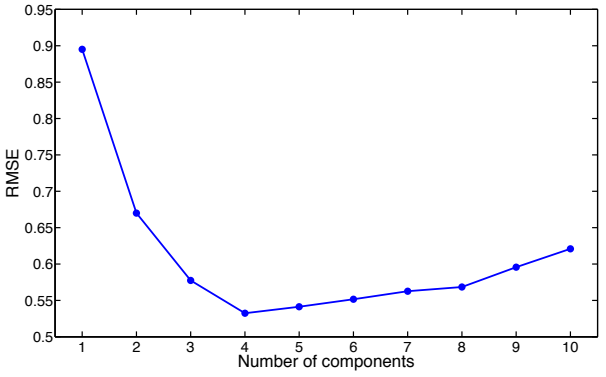


**Fig. 7.** Modelling error as a function of the number of latent variables (model order)

measurement of the reaction rate, assuming the heat removed from the reactor through the cooling spiral remains constant.

### 4.3 Model Order Determination

To select the optimal number of latent variables $R$ for the MPLS model, the number of latent variables (i.e., the model order) is varied, and the influence on the performance is observed. To ensure the best model accuracy, leave-one-out (LOO) crossvalidation is used instead of 10-fold crossvalidation.

As depicted in Figure 7, the performance initially increases when the model order increased. If the model order is increased further, the performance curve passes through a shallow optimum, after which the performance degrades as overfitting occurs. Based on this graph, it is clear that adding extra latent variables beyond the fourth has a negative impact on the model performance gain. Hence, four latent variables are used for the inferential MPLS sensor.

### 4.4 Model Training

Before the model is identified, the performance of the model structure with three input variables and four latent variables is investigated. Figure 8 depicts the LOO crossvalidation predictions for the training batches (RMSE equals 0.532). It is clear that model predictions and laboratory measurements of the polymer viscosity show good agreement. It is therefore concluded that the identified MPLS model structure makes accurate predictions of the final polymer viscosity.

Next, the final full model is identified by training on all 72 available batches. The model captures 82.3% of the variance in the quality variable $\mathbf{Y}$.

Outliers in the training data set are identified via the $T^2$ and $Q^2$ statistics. The $T^2$ analyses the similarity between the batches. A large value indicates that a batch is different from the batches from the *normal operating conditions*.

$$T_i^2 = \mathbf{T}_i \, \Sigma_{\mathbf{T}}^{-1} \mathbf{T}_i^T \tag{6}$$

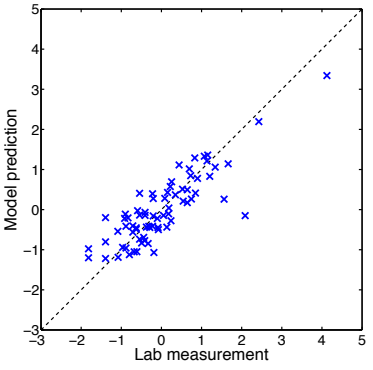


**Fig. 8.** Comparison of model predictions and lab measurements of the viscosity

The matrix $\Sigma_{\mathbf{T}}$ is the covariance matrix of the scores $\mathbf{T}$ from the training batches and is obtained via LOO crossvalidation. $T^2$ is $F_{(R,I_{\mathrm{tr}}-R)}$-distributed; its upper control limit $u_T$ at a specified tolerance level $\alpha$ is given by [4,19,20,25]

$$u_T = \frac{R(I_{\mathrm{tr}}^2 - 1)}{I_{\mathrm{tr}}(I_{\mathrm{tr}} - R)} \; F\left(R, I_{\mathrm{tr}} - R; \alpha\right). \tag{7}$$

Here, $I_{\mathrm{tr}}$ is the total number of training batches, and $F\left(R, I_{\mathrm{tr}} - R; \alpha\right)$ is the upper critical value of the $F$-distribution with $R$ numerator degrees of freedom and $I_{\mathrm{tr}} - R$ denominator degrees of freedom, and tolerance $\alpha$.

The $T^2$-statistic for the training batches is depicted in Figure 9. From this plot, it can be seen that only one batch exceeds the 99% confidence value, and is identified as abnormal by the MPLS model. This batch corresponds with the highest observed viscosity value in the training data. Because the correct prediction of this too high viscosity value is preferred over a simple identification as *off-spec*, this outlying batch is nonetheless retained in the training data set.

The $Q^2$ indicates how well the MPLS model fits each batch by analyzing the residuals $\mathbf{E_X}$ for each batch $i$: a large $Q^2$ indicates the model is invalid.

$$Q_i^2 = \mathbf{E_X}\mathbf{E_X}^T = \left(\mathbf{X}_i - \mathbf{T}_i\mathbf{P}^T\right)\left(\mathbf{X}_i - \mathbf{T}_i\mathbf{P}^T\right)^T \tag{8}$$

The $Q^2$ statistic follows a $g\chi_k^2$ distribution, where $g$ and $k$ are determined via LOO crossvalidation. The control limit with tolerance $\alpha$ is [4,19,20,25]

$$u_Q = \frac{\sigma_Q^2}{2\mu_Q} \; \chi^2\left(\frac{2\mu_Q^2}{\sigma_Q^2}; \alpha\right) \tag{9}$$

where $\mu_Q$ and $\sigma_Q$ are the mean and standard deviation of the LOO $Q^2$ statistics for the training batches, and $\chi^2\left(2\mu_Q^2/\sigma_Q^2; \alpha\right)$ is the upper critical value of the $\chi^2$ distribution with $2\mu_Q^2/\sigma_Q^2$ degrees of freedom.

The $Q^2$ for all training batches is located well below the control limit with 99% confidence. This indicates that the MPLS model is a good fit for all batches.

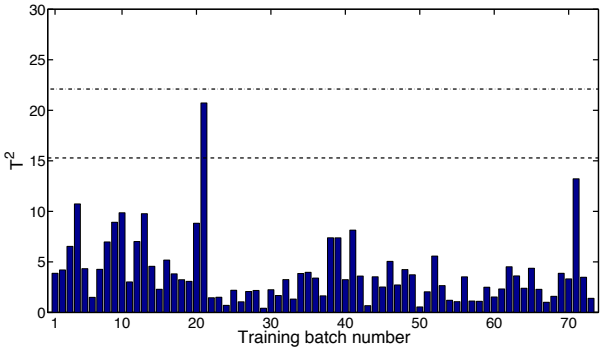


**Fig. 9.** $T^2$-statistic for the training batches. The horizontal lines indicate the 99% (– –) and 99.9% (–·–) confidence limits.

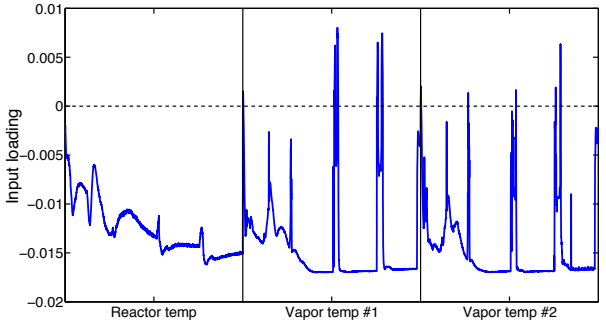


**Fig. 10.** MPLS model loadings **P** for all variables for the first principal component

The loadings **P** of the MPLS model are also investigated. Because the columns of the input data matrix **X** are not independent but contain profiles of only three variables, the elements of the loadings matrix **P** also represent profiles. Figure 10 depicts the loading profiles for each of the three input variables.

The negative loading values for the reactor temperature indicate a negative correlation between the reactor temperature and the polymer's viscosity. This negative correlation is explained physically: an increase in the reactor temperature leads to a higher polymerization rate, when monomer concentrations are equal. The higher reaction rate causes a decrease in overall polymer chain length, which in turn causes a decrease in the polymer's viscosity. For the two vapor temperature measurements, a similar negative correlation with the final viscosity is observed, and identical conclusions can be drawn.

During the initiator shots, however, a positive –or less negative– correlation between the vapor temperatures and the viscosity is observed. At these times, the reaction heat used for evaporation is instead used for activating the initiator. Hence, the vapor phase cools down while the reaction rate increases, again leading to a lower viscosity. At the same time, a sudden increase in the negative

correlation can also be seen for the reactor temperature. A higher reactor temperature means a quicker activation of the added initiator, again leading to a higher polymerization rate and the corresponding decrease in polymer viscosity.

Finally, the training performance of the full model is studied by comparing the model predictions with the lab measurements. The graphical comparison yields a plot similar to Figure 8.

## 5 Intermediate Model Identification

The MPLS model identified in Section 4 takes completed profiles as inputs. Therefore, it is only capable of predicting the polymer viscosity after completion of the batch. This prediction can be made as soon as the cooling of the batch starts, well before lab analysis results are obtained. However, in order to better control batch operation, model predictions must be available during the batch run. While different techniques are available for making MPLS predictions using only partially known inputs, as is the case during batch operation, these methods all assume aligned profiles are available [2,7].

Because (HD)DTW requires the final point of the new profile to be known to determine the warping path. Hence, it can only be applied once a batch phase is completed. Although an online implementation of HDDTW is presented in [9], an intermediate model approach is adopted in this work to minimize the online computational requirements. A first intermediate model takes only the data from the first and second batch phase as input, and can be used to make a viscosity prediction at approximately $t_f/2$ (50% batch completion). A second model takes the first three phases as inputs, making a prediction at 75% completion.

Using the same procedure as detailed in Sections 4.2 and 4.3, the optimal input variables and number of latent variables for each of the intermediate MPLS models are identified. The optimal input variables are identical to those obtained in Section 4.2, while the optimal number of latent variables increases to five.

Next, the crossvalidation performance of these intermediate models is studied. Table 1 compares the average crossvalidation RMSE values of the intermediate models with those of the full model. From this table, it is clear that the two intermediate models exhibit a performance similar to the full model. This implies that accurate viscosity estimations can be obtained already at $t_f/2$, only halfway throughout the batch process.

**Table 1.** Crossvalidation and validation performance comparison of the full and partial models

| | | RMSE | |
|---|---|---|---|
| | Input phases | Crossvalidation | Validation |
| Intermediate model #1 | 1-2 | 0.549 | 0.621 |
| Intermediate model #2 | 1-2-3 | 0.557 | 0.572 |
| Full model | 1-2-3-4 | 0.532 | 0.568 |

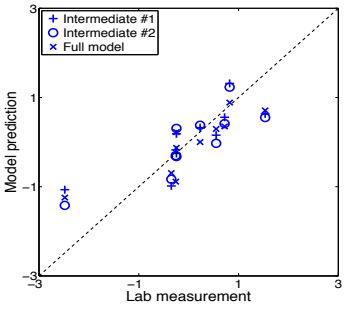


**Fig. 11.** Comparison of model predictions and laboratory measurements of the polymer viscosity for the full model and both intermediate models.

Finally, the final model weights for the intermediate models are identified by training on all available batches.

## 6 Validation Results

Finally, the three MPLS models identified in Sections 4 and 5 are validated on 10 additional batches. The comparison between the model predictions and lab viscosity measurements is listed in Table 1 and depicted in Figure 11.

All validation batches have $T^2$ and $Q^2$ values below the the 99% control limit. Hence, all model predictions of the polymer viscosity are considered reliable.

This result might suggest that the final 25% or even 50% of the batch operation has no significant impact on the final polymer viscosity, and could be removed from the data set without compromising the model performance. However, this observation is only valid if the final batch phases exhibit no abnormal behavior. Hence, by including the second half of the batch in the second intermediate model and the full model, abnormal operation of the chemical batch reactor can be detected.

It is clear that all three models give accurate viscosity predictions, in line with the observations of Section 5. Hence, it is concluded that all three models perform equally well.

## 7 Conclusions

In this paper, an inferential sensor capable of predicting the final viscosity of a polymer produced in a chemical batch reactor was identified, to alleviate the problem of difficult batch process control. By enabling the early prediction of final product quality, corrective actions can be taken during the batch run, resulting in better quality control and the production of less off-spec batches.

A *hybrid dynamic time warping* algorithm was implemented to bring all batches to equal length and to align significant events. The algorithm yielded

very good results and was robust with respect to measurement noise. The optimal input and model order of the partial least squares model were selected during the identification of the full batch model. In addition, intermediate models were developed to predict the final product quality after 50% or 75% completion. The accuracy of these intermediate models was equal to that of the full model. Finally, the performance of the three models was validated on extra industrial data. Again, all models yielded similar prediction qualities, comparable with the results obtained during training.

This leads to the conclusion that the developed inferential sensor is indeed capable of making accurate predictions of the final polymer viscosity well before the end of the batch run. Because of this valuable result, it is possible to exploit the estimation provided by the sensor to control the batch, resulting in fewer off-spec (wasted) batches. With the results obtained in this work, the financial losses associated with an off-spec batch are reduced by 30%.

While the inferential sensor developed in this work yields accurate final polymer viscosity predictions, estimations are not available online. Hence, control of the batch remains difficult. Therefore, future work will consist of the implementation of an online HDDTW algorithm. This will allow the estimation of the final product viscosity at more regular intervals, enabling the operator to actively monitor a running batch, and adjust process parameters to control the final product viscosity.

**Acknowledgements.** Work supported in part by Projects OT/09/25/TBA PVF/10/002 (OPTEC Optimization in Engineering Center), Project KP/09/005 (SCORES4CHEM) and IUAP P6/04 (DYSCO). J. Van Impe holds the chair Safety Engineering sponsored by the Belgian chemistry and life sciences federation essenscia. Scientific responsibility is assumed by its authors.

## References

1. Aach, J., Church, G.: Aligning gene expression time series with time warping algorithms. Bioinformatics 17, 495–508 (2001)
2. Arteaga, F., Ferrer, A.: Dealing with missing data in MSPC: several methods, different interpretations, some examples. J. Chemometr. 16, 408–418 (2002)
3. Caiani, E.G., Porta, A., Baselli, G., Turiel, M., Muzzupappa, S., Pieruzzi, F., Crema, C., Malliani, A., Cerutti, S.: Warped-average template technique to track on a cycle-by-cycle basis the cardiac filling phases on left ventricular volume. IEEE Computers in Cardiology 25, 73–76 (1998)
4. Choi, S.W., Martin, E.B., Morris, A.J., Lee, I.-B.: Dynamic model-based batch process monitoring. Chem. Eng. Sci (2007), doi:10.1016/j.ces.2007.09.046
5. Dorsey, A.W., Lee, J.H.: Building inferential prediction models of batch processes using subspace identification. J. Proc. Contr. 13, 397–406 (2003)
6. García-Munoz, S., Kourti, T., MacGregor, J.F., Mateos, A.G., Murphy, G.: Troubleshooting of an industrial batch process using multivariate methods. Ind. Eng. Chem. Res. 42, 3592–3601 (2003)
7. García-Munoz, S., Kourti, T., MacGregor, J.F.: Model predictive monitoring for batch processes. Ind. Eng. Chem. Res 43, 5929–5941 (2004)

8. Geladi, P., Kowalski, B.R.: Partial least-squares regression: a tutorial. Anal. Chim. Acta 185, 1–17 (1986)
9. Gins, G.: Modelling of (bio)chemical processes using data-driven techniques, PhD Thesis, Faculteit Ingenieurswetenschappen, Katholieke Universiteit Leuven, Belgium (2007)
10. Itakura, F.: Minimum prediction residual principle applied to speech recognition. IEEE Trans. on Acoustics, Speech and Signal Proc. ASSP-23, 52–57 (1975)
11. Kassidas, A., MacGregor, J.F., Taylor, P.A.: Synchronization of batch trajectories using dynamic time warping. AIChE J. 44(4), 864–875 (1998)
12. Keogh, E.J., Pazzani, M.J.: Derivative dynamic time warping. In: First SIAM International Conference on Data Mining, Chicago, IL, 2001 (2001)
13. Kourti, T., Nomikos, P., MacGregor, J.F.: Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS. J. Proc. Contr. 5, 277–284 (1995)
14. Kourti, T.: Multivariate dynamic data modeling for analysis and statistical process control of batch processes, start-ups and grade transitions. J. Chemometr. 17, 93–109 (2003)
15. Lee, J.H., Dorsey, A.W.: Monitoring of batch processes through state-space models. AIChE J. 50(6), 1198–1210 (2004)
16. Lu, N., Yao, Y., Gao, F.: Two-dimensional dynamic pca for batch process monitoring. AIChE J. 51(12), 3300–3304 (2005)
17. McCready, C.: Model predictive multivariate control. In: Proc. 2nd European Conference on Process Analytics and Control Technology, p. 82 (2011)
18. Nomikos, P., MacGregor, J.F.: Monitoring of batch processes using multi-way principal component analysis. AIChE J. 40(8), 1361–1375 (1994)
19. Nomikos, P., MacGregor, J.F.: Multivariate SPC charts for monitoring batch processes. Technometr. 37(1), 41–59 (1995)
20. Nomikos, P., MacGregor, J.F.: Multiway partial least squares in monitoring batch processes. Chemometr. Intell. Lab. Syst. 30, 97–108 (1995)
21. Ramaker, H.-J., Van Sprang, E.N.M., Westerhuis, J.A., Smilde, A.K.: Dynamic time warping of spectroscopic BATCH data. Anal. Chim. Acta 498, 133–153 (2003)
22. Ratanamahatana, C.A., Keogh, E.J.: Three myths about dynamic time warping. In: Proceedings of SIAM International Conference on Data Mining (SDM 2005), Newport Beach, California, USA, pp. 506–510 (2005)
23. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. on Acoustics, Speech, and Signal Proc. ASSP-26(1), 43–49 (1978)
24. Wold, S., Geladi, P., Ebensen, K., Öhman, J.: Multi-way principal components-and PLS-analysis. J. Chemometr. 1(1), 41–56 (1987)
25. Tracy, N., Young, J., Mason, R.: Multivariate control charts for individual observations. J. Qual. Technol. 24(2), 88–95 (1992)