

Article

Polymethyl Methacrylate Quality Modeling with Missing Data Using Subspace Based Model Identification

Nikesh Patel, Kavitha Sivanathan and Prashant Mhaskar * 

Department of Chemical Engineering, McMaster University, Hamilton, ON L8S 4L7, Canada; patelna@mcmaster.ca (N.P.); sivank1@mcmaster.ca (K.S.)

* Correspondence: mhaskar@mcmaster.ca

Abstract: This paper addresses the problem of quality modeling in polymethyl methacrylate (PMMA) production. The key challenge is handling the large amounts of missing quality measurements in each batch due to the time and cost sensitive nature of the measurements. To this end, a missing data subspace algorithm that adapts nonlinear iterative partial least squares (NIPALS) algorithms from both partial least squares (PLS) and principal component analysis (PCA) is utilized to build a data driven dynamic model. The use of NIPALS algorithms allows for the correlation structure of the input–output data to minimize the impact of the large amounts of missing quality measurements. These techniques are utilized in a simulated case study to successfully model the PMMA process in particular, and demonstrate the efficacy of the algorithm to handle the quality prediction problem in general.

Keywords: polymethyl methacrylate; subspace identification; data driven model identification; missing data



Citation: Patel, N.; Sivanathan, K.; Mhaskar, P. Polymethyl Methacrylate Quality Modeling with Missing Data Using Subspace Based Model Identification. *Processes* **2021**, *9*, 1691. <https://doi.org/10.3390/pr9101691>

Academic Editor: Jie Zhang

Received: 23 July 2021

Accepted: 17 September 2021

Published: 22 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Polymethyl methacrylate (PMMA) is an important industrial polymer with widespread applications ranging from plastics to electronics. PMMA production occurs via a specific recipe in a batch process in order to ensure the optimal product quality. The key parameter to determining product quality is the molecular weight distribution of PMMA. However, measuring the entire molecular weight is a long, complex procedure; therefore, the number and weight average molecular weights are used as outcome measures instead. These quality variables are measured at the end of the batch and, as such, must be controlled tightly.

In order to achieve good quality control of the PMMA process, it is important to develop a process model capable of handling batch data. Batch processes are unique in that they tend towards the low volume production of high value products, as this allows for poor quality batches to be discarded [1]. Each discarded batch represents a significant loss in revenue, motivating the need for advanced batch control approaches and, fundamentally, the development of accurate quality models. Recent advances in computing technology have led to increased amounts of historical data being available, making data-driven modeling a viable choice for model identification [2–7]. However, these techniques still have challenges when dealing with nonlinear processes and scenarios where data is missing, as is the case with quality variables.

An important consideration when identifying data-driven models is the choice of input and output variables. While the distinction is typically based on variables that are controlled in comparison to those that are measured, the batch process also introduces a separate type of output variable: quality variables. Quality variables are still considered to be outputs from the process, however, they are not measured continuously, nor are they always measured online. Quality measurements are often calculated based on the regular measured outputs or are determined by separate analyses of the batch. This difference is an important consideration for data driven modeling techniques, as one assumption that

is prevalent in process modeling scenarios is that the inputs and outputs are sampled at a single and uniform sampling rate. In practice, industrial processes often have different sampling rates for input and output variables. Additionally, processes record quality measurements at an even slower rate compared to traditional inputs and outputs. This leads to scenarios where inputs and outputs have some missing values due to differences in sampling rates, whereas quality measurements are available at extremely low frequencies. This presents a challenge to traditional data driven modeling approaches that require complete data sets to identify a model.

When attempting to model industrial batch processes using data containing missing observations, a common data-driven technique used is partial least squares (PLS) [8]. In this technique, process data from multiple batches is taken and projected into two lower dimensional subspaces (latent variable space). This ensures that the relationship between the correlated input and output space variables is maintained and is characterized by the independent latent variables [9]. PLS techniques are capable of handling missing data since they utilize the covariant structure between the input and output variables from the original variable space. This inherent ability is one of the key properties that makes PLS techniques suitable for modeling batch processes. The application of PLS techniques to batch process modeling has been previously explored and one successful approach is to utilize batchwise unfolding of the data [9–11]. Process data from each batch is unfolded into a single PLS observation that is subsequently related to quality variables. This approach can be applied to on-line process data by utilizing data imputation techniques to make predictions on the missing data observations. While this approach has been well-documented in handling industrial batch data, it requires batch alignment in order to account for varying batch duration. Furthermore, the approach identifies a time-varying model and, since PLS inherently does not distinguish between input and output variables, it has limited applicability in traditional model predictive control and scenarios where batch duration is a decision variable.

Another technique that is suitable for building process models is subspace identification [5–7,12], which has been appropriately modified for handling batch data using Hankel matrices [1]. Note that subspace identification is different from PLS because it explicitly distinguishes between input and output variables [5–7,12]. Subspace identification consists of two distinct steps: identifying a state trajectory from historical input and output batch data, and determining the system matrices using a least squares solution. To achieve these outcomes, subspace identification utilizes a range of techniques from canonical variate analysis [13,14], numerical algorithms [15] and multivariate output error state space algorithms [16]. One common technique to subspace identification is singular value decomposition (SVD) of the matrices [5,6]. However, SVD requires matrices to be full-rank, making it unsuitable for handling batch data with missing observations [17].

One way to handle the problem of missing data with subspace identification is through a lifting technique [18]. The prevalent assumption is that all the inputs are sampled at one frequency and all the outputs are sampled at another frequency. A model can then be derived using the grouped inputs and outputs wherever the sampling frequencies align. This process involves computed control moves over the larger sampling interval instead of every sampling instance. Lifting has many advantages, including the ability to convert a multi-rate system into a single rate system, but the key benefits lie in model predictive control. However, this process is unsuited for the quality control problem, as quality measurements are not frequent enough to provide sufficient excitation for modeling. These measurements are only available at percentages less than 10%, making a single rate system impractical. Thus, a modeling approach that can use the lower frequency sampling rate of the inputs and outputs is required. The previous work [19] developed a technique for handling missing data in subspace identification approach with a focus on solving the problem of random missing data. The present work recognizes that the missing data subspace approach can be applied to the broader quality measurement induced missing data problem. Thus, the present manuscript provides a solution for instances where the

low frequency of quality measurements in relation to other online measurements results in a missing data problem (as in the simulated PMMA process example).

To overcome these challenges and model the PMMA process, the current paper focuses on a recent modification to subspace identification [20] that treats nonuniform sampling rate data as a missing data problem. Specifically, the addition of PCA and PLS steps to the subspace identification approach permits accounting for the missing quality measurements. While the use of PCA and PLS techniques is not a novel introduction to model identification (see [21,22]), the reduced latent variable space is marginally affected by missing data, thus allowing for the treatment of nonuniform sampling problems as a case of randomly missing data. The first step is to use latent variable methods (PCA followed by PLS) to identify a reduced dimensional space for the variables which accounts for missing observations. The second step replaces SVD with PCA to identify the states of the system, whereupon traditional subspace approaches can be utilized. The rest of the paper is organized as follows: Section Preliminaries presents the Polymethyl Methacrylate (PMMA) process, followed by a review of traditional subspace identification. The next subsection presents the Subspace Identification approach that can readily handle missing data issues. In the results section, identification and validation results for the PMMA process are presented. Finally, concluding remarks are made.

2. Preliminaries

2.1. Polymethyl Methacrylate (PMMA) Process Description

PMMA is produced via free-radical polymerization in a jacketed reactor with a continuous stirrer. At the start of each batch, the reactor is charged with methyl methacrylate (monomer), AIBN (initiator), and toluene (solvent). The system is controlled by manipulating the heat input to the reactor, which is based on the coolant/heating fluid flow rate. Table 1 presents the nominal set of batch conditions used to initialize the PMMA process with standard deviation of σ^2 . In order to provide good quality batches, a nominal temperature trajectory known to yield good quality results is tracked for each batch. Proportional-integral (PI) tracking is used to track the batch temperature by manipulating the jacket temperature. The batch is then terminated after the polymer density reaches the desired set-point.

Table 1. Initial batch conditions for PMMA reaction.

| Variable | Value | σ^2 | Units |
|-------------------------|-----------------------|----------------------|-------------------|
| Initiator Concentration | 2.06×10^{-2} | 1.0×10^{-3} | kg/m ³ |
| Monomer Concentration | 4.57 | 0.5 | kg/m ³ |
| Temperature | 61 | 2.5 | °C |

A first principles dynamic model is utilized as a test bed consisting of nine states governed by a set of nonlinear algebraic and differential equations. The nine system states are: reactor temperature, monomer and initiator concentrations and six molecular weight distribution moments. This dynamic PMMA model has been taken from [23] with suitable adaptations [24,25] included [26]. To replicate natural process variability in each batch, the initial conditions of the PMMA process are varied using a normal distribution around the nominal conditions as shown in Table 1. Three variables are assumed to be measured continuously: reactor temperature, reactant volume and stirrer torque. The volume measurements are used to calculate the density, which is used to measure the extent of the reaction, allowing the batch to be terminated once the reaction has reached a certain threshold, as opposed to terminating at a fixed batch length. This leads to variable batch lengths depending on the time required to reach the target density.

2.2. Subspace Identification

This subsection provides a brief description of existing subspace identification methods that provide the basis for the PCA and PLS based modifications. These techniques utilize input/output data to identify a state space model (typically) around steady-state operating conditions. Specifically, given a batch of s measurements of the inputs $u_k \in \mathbb{R}^p$ and outputs $y_k \in \mathbb{R}^q$, a system with order n can be determined in the following form:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \quad (1)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k, \quad (2)$$

where the identification approach identified the system order n along with the system matrices $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times p}$, $\mathbf{C} \in \mathbb{R}^{q \times n}$, $\mathbf{D} \in \mathbb{R}^{q \times p}$ (up to within a similarity transformation). Considerable work has been done to show how this model can be developed for batch processes. Thus, consider k as the sampling instant from the start of the batch and b denoting the batch number [1]. The identified state trajectory matrix for each batch is represented as:

$$\mathbf{X}_{i+1}^{(b)} = \begin{bmatrix} \mathbf{x}^{(b)}[i+1] & \cdots & \mathbf{x}^{(b)}[i+j^{(b)}] \end{bmatrix} \quad \forall b = 1, \dots, nb \quad (3)$$

$$\mathbf{X}_{i+1} = \begin{bmatrix} \mathbf{X}_{i+1}^{(1)} & \mathbf{X}_{i+1}^{(2)} & \cdots & \mathbf{X}_{i+1}^{(nb)} \end{bmatrix} \quad (4)$$

where nb is the total number of batches used for identification.

Using the state trajectory matrix, the system matrices can be estimated easily using the following matrix regression steps:

$$\mathbf{Y}_{reg}^{(b)} = \begin{bmatrix} \mathbf{x}^{(b)}[i+2] & \cdots & \mathbf{x}^{(b)}[i+j^{(b)}] \\ \mathbf{y}^{(b)}[i+1] & \cdots & \mathbf{y}^{(b)}[i+j^{(b)}-1] \end{bmatrix} \quad (5)$$

$$\mathbf{X}_{reg}^{(b)} = \begin{bmatrix} \mathbf{x}^{(b)}[i+1] & \cdots & \mathbf{x}^{(b)}[i+j^{(b)}-1] \\ \mathbf{u}^{(b)}[i+1] & \cdots & \mathbf{u}^{(b)}[i+j^{(b)}-1] \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} \mathbf{Y}_{reg}^{(1)} & \cdots & \mathbf{Y}_{reg}^{(nb)} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{reg}^{(1)} & \cdots & \mathbf{X}_{reg}^{(nb)} \end{bmatrix} \quad (7)$$

2.3. Noniterative Partial Least Squares (NIPALS) Algorithm

This subsection provides a brief description of the existing NIPALS based PCA and PLS subspace identification approach. These steps are used to carry out regressions in the traditional subspace approach that rely on full rank matrices, since PCA and PLS can handle missing data. This approach is fully described in Patel et al. [19].

The first step in subspace identification is to project the future outputs and past inputs and outputs perpendicular to future inputs to remove the correlation between the past data and future inputs. This can be achieved using a NIPALS algorithm for PCA. The PCA step removes all of the variance associated with the future inputs, and the remaining correlations are the result of the effect of current states on future outputs. The next step in the NIPALS approach is to identify the relationship between the past inputs and outputs and the future outputs. This relationship provides the basis of subspace identification techniques, as it relates directly to the underlying states of the process. In order to identify this relationship, a NIPALS algorithm for PLS is applied. Here, the two blocks are chosen such that the first block contains the past inputs and outputs and the second block contains the future outputs. Finally, after carrying out PLS, the first block contains only known values (coefficients from the past inputs and outputs). This matrix can then be appropriately decomposed to matrices to match the second block, resulting in a valid state space representation. In traditional subspace, the relationship is then identified through singular value decomposition (SVD). However, SVD is not suitable for handling matrices with missing data, and PCA can be

used instead. It then follows that the scores from PCA are a valid realization of the process states and can then be used in Equation (7) to compute the system matrices (\mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{D}) via linear regression.

Remark 1. *This approach is particularly important when considering the quality control problem, and opens up the possibility of direct control of quality. Thus, since the model can be utilized to predict quality variables during the intervals which it is not measured, a classical control structure such as PID can be utilized to directly control the (predicted) quality to a desired trajectory, terminating at the desired final quality.*

3. PMMA Model Identification

Two case studies are considered for the PMMA process, where the output data is lost in storage (thus, the PI controller is able to run in the closed-loop using the available temperature measurements. Note that the approach and implementation readily holds for the situation where the measurements would simply be unavailable at random times, and would need a change in how the PI controller is implemented). Case 1 contains training data with 30% random missing output data and quality data measurements retained at intervals of 1 in every 10, 20 and 30 time steps, respectively. Case 2 contains training data with quality measurement available every 10 time steps and random missing output data ranging from 5% up to 30%. Thus, the effects of both lower frequency of quality measurement availability and increasing missing data are independently investigated.

Remark 2. *The key difference between data imputation approaches and the missing data approach is that imputation can lead to inconsistency between the imputed data (in terms of the model structure implied) and the model structure being identified. Thus, in the comparison case study, where linear interpolation techniques are utilized, it is seen that linear interpolation is not consistent with the type of model ultimately being identified (linear time invariant dynamic model), resulting in the inability to capture the process trends sufficiently, leading to poor model performance. The success of the approach is based on the format of the data-blocks, which guarantees a high degree of covariance due to batch unfolding of the Hankel matrices and the temporal nature of the data.*

3.1. Model Identification

For each case, 15 training batches under proportional integral (PI) control are generated from the PMMA process and are used to identify two subspace models. To compare the present approach against existing results, first a model is identified using traditional subspace identification with linear interpolation of the measured outputs to determine missing values in the training batches and, subsequently, the proposed approach using PCA/PLS techniques to handle infrequent quality sampling is implemented. The two models are then validated against a new set of 15 batches (under PI control) that are also modified to have missing data similar to the training batches to determine the validation error.

Remark 3. *Note that the number of states and lags are a user specified parameter. The choice is based on the best training fit defined as the lowest error between the predictions and the process outputs. The states and lags used are specific to the present implementation and future work will explore further the relationship between the number of states and model quality.*

3.2. State Observer

After identifying the state space model an important consideration before utilizing it for online predictions is to ensure that the model states converge with the process. This approach utilizes the Luenberger observer, which updates the states using feedback from the outputs weighted using a gain matrix \mathbf{L} in the following method:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{L}(y_{process} - \hat{y}) \quad (8)$$

To handle missing outputs when the model is run with a new batch, this approach uses linear interpolation until the process outputs converge within a tolerance of 0.1% or until 40 min have passed (the sampling instance when this occurs is denoted by c). After that, the model is run online and is used to predict the process outputs. The model's effectiveness can then be determined based on the prediction error until the end of the batch. The error is determined by the sum of the difference between each available process output ($y_{process}$) and model prediction (\hat{y}) for the duration that the models are used for prediction (i.e., after the Luenberger observer has converged).

This value is then normalized by the number of process outputs available to account for differences in batch duration as well as missing observations. The validation error for each batch is calculated as:

$$\text{Validation Error} = \frac{\sum_{k=c}^s |\hat{y}(k) - y_{process}(k)|}{\# \text{ of predictions}} \quad (9)$$

where k is the batch index, c is the index when the Luenberger observer converges and s is the batch length.

3.3. Case 1: Missing Quality Data

In this case, the amount of missing quality data was varied to retain quality measurements in intervals of 10, 20 and 30 sampling times, resulting in missing data percentages of 90%, 95% and 97%, respectively. The reason for doing this is to test the utility of the approach for situations where the quality measurements may take longer times to be obtained. Additionally, output data from the training batches are randomly deleted to obtain a missing data percentage of 30%. Both the missing data approach and the traditional subspace approach were then used to develop a subspace model. To handle the missing measurements, the traditional subspace identification approach uses linear interpolation in order to have a full rank matrix. The two models are then validated against validation batches where both output and quality data entries were removed, similar to training batches.

Table 2 shows the average validation error of the traditional subspace model and missing data approach from modeling 15 validation batches. The missing data approach consistently has a lower error compared to the traditional subspace model. Figure 1 shows the predictions for reactant conversion from the proposed approach, traditional approach, as well as the output estimates using the Luenberger observer for the first 40 time-steps from both models compared to the true process for one batch. Figures 2 and 3 show the predictions for the number average molecular weight and weight average molecular weight from both models compared to the true process. Looking at each of the model predictions, the three figures clearly show how the missing data approach is able to more accurately model the true process in comparison to the traditional subspace identification.

Table 2. Average validation error for two models with missing quality data and 30% random output data.

| Model | Keep 10 90% Missing | Keep 20 95% Missing | Keep 30 97% Missing |
|-------------|---------------------|---------------------|---------------------|
| Traditional | 56,106 | 73,996 | 61,190 |
| Proposed | 10,149 | 8779 | 9781 |
| Lags | 9 | 12 | 14 |

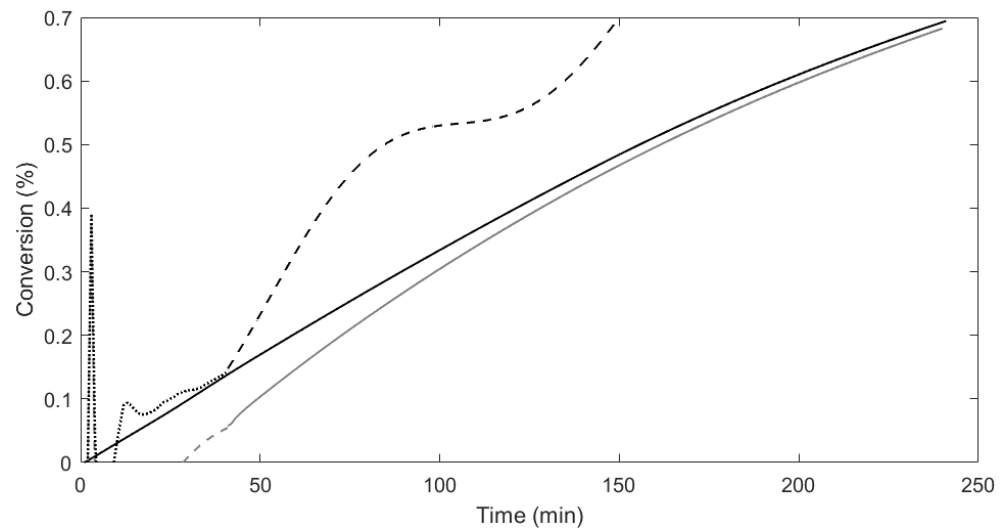


Figure 1. The conversion predictions from both models with every tenth quality measurement retained and 30% missing output data. The dashed grey lines represent Luenberger observer estimates until the 40th minute and the solid grey line shows the predictions from the proposed approach in comparison to the process, represented by the solid black line. The dotted and dashed black lines show the predictions made by the Luenberger observer and traditional subspace identification using linear interpolation, respectively.

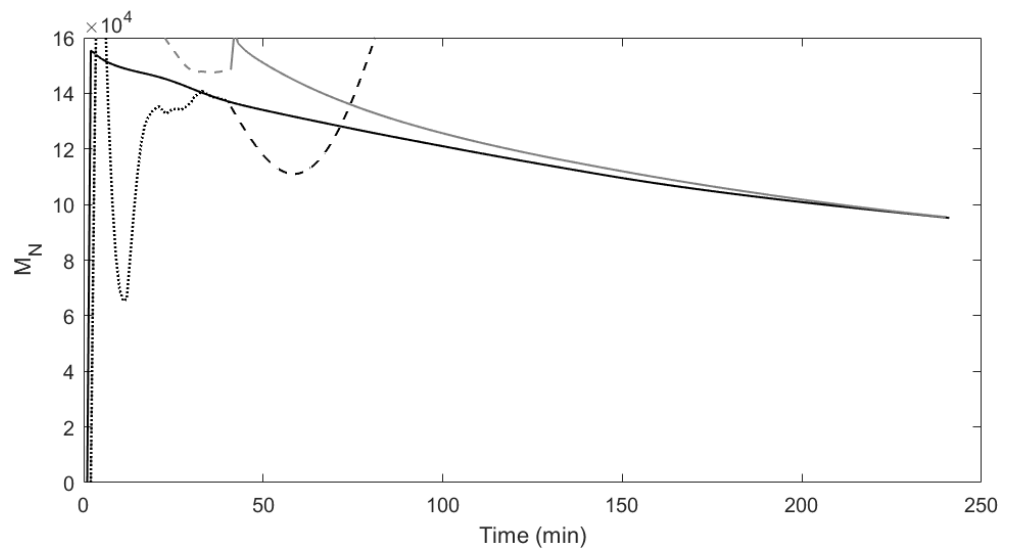


Figure 2. The number average molecular weight predictions from both models with every tenth quality measurement retained and 30% missing output data. The dashed grey lines represent estimates made by the Luenberger observer until the 40th minute and the solid grey line shows the predictions from the proposed approach in comparison to the process, represented by the solid black line. The dotted and dashed black lines show the estimates made by the Luenberger observer and traditional subspace identification using linear interpolation, respectively.

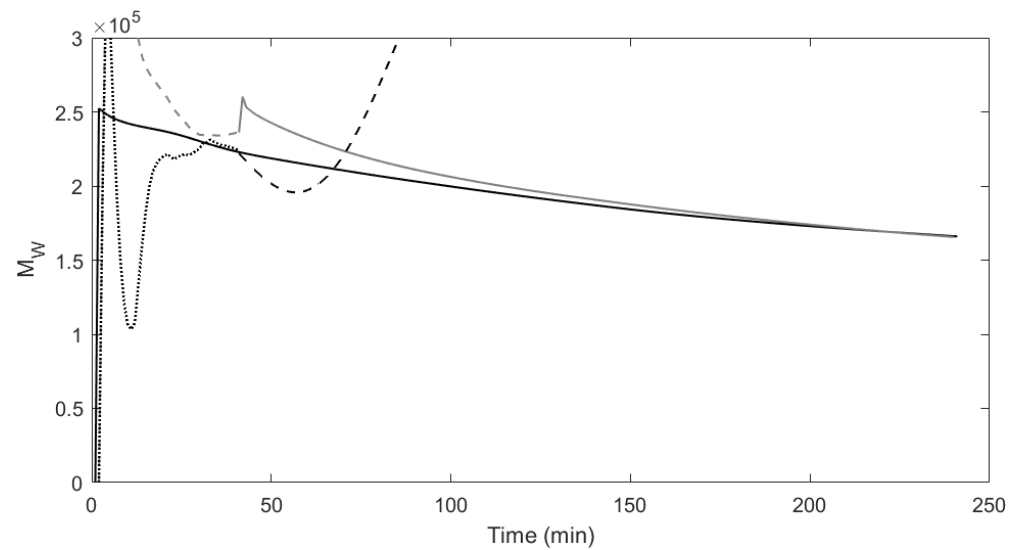


Figure 3. The weight average molecular weight predictions from both models with every tenth quality measurement retained and 30% missing output data. The dashed grey lines represent estimates made by the Luenberger observer until the 40th minute and the solid grey line shows the predictions from the proposed approach in comparison to the process, represented by the solid black line. The dotted and dashed black lines show the estimates made by the Luenberger observer and traditional subspace identification using linear interpolation, respectively.

3.4. Case 2: Random Missing Output Data

In this case, the output data from training batches is randomly deleted in order to achieve missing data percentages from 5% up to 30% while every tenth quality variable measurement is kept (90% missing outputs). The missing data approach is then compared against traditional subspace identification with linear interpolation. The models are then validated against validation batches with both output and quality data entries removed in a similar manner to training batches.

Table 3 shows the average validation error in predicting quality variable data for 15 batches from the traditional subspace identification and missing data approach. In comparison to the traditional subspace approach, the missing data approach is able to consistently make more accurate predictions. Figure 4 shows the conversion predictions made by the proposed approach, traditional subspace approach, as well as the predictions from the Luenberger observer for the first 40 time-steps for both models compared to the true process for one validation batch. Figures 5 and 6 show the predictions for number average molecular weight and weight average molecular weight from both models compared to the true process. The three graphs clearly show that the proposed approach is able to more accurately model and predict the process data compared to traditional subspace identification.

Table 3. Average validation error for two models with every tenth quality measurement retained and random missing output data.

| Model | Error 5% | Error 10% | Error 15% | Error 20% | Error 25% | Error 30% |
|-------------|----------|-----------|-----------|-----------|-----------|-----------|
| Traditional | 28,649 | 29,760 | 34,187 | 36,041 | 36,645 | 56,106 |
| Proposed | 12,674 | 8016 | 9342 | 10,450 | 10,242 | 10,149 |

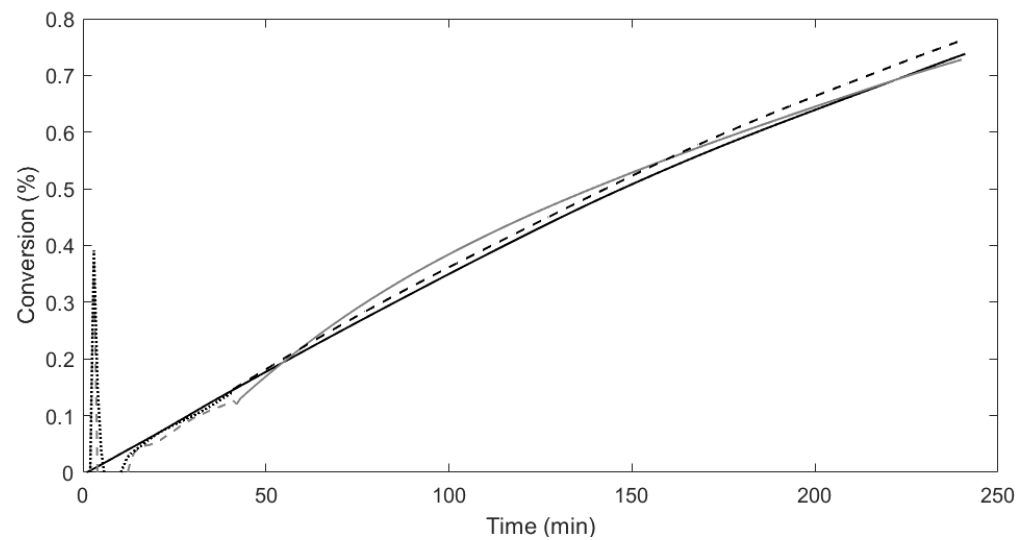


Figure 4. The conversion predictions from both models with every tenth quality measurement retained and 20% missing output data. The dashed grey lines represent estimates made by the Luenberger observer until the 40th minute and the solid grey line shows the predictions from the proposed approach in comparison to the process, represented by the solid black line. The dotted and dashed black lines show the estimates made by the Luenberger observer and traditional subspace identification using linear interpolation, respectively.

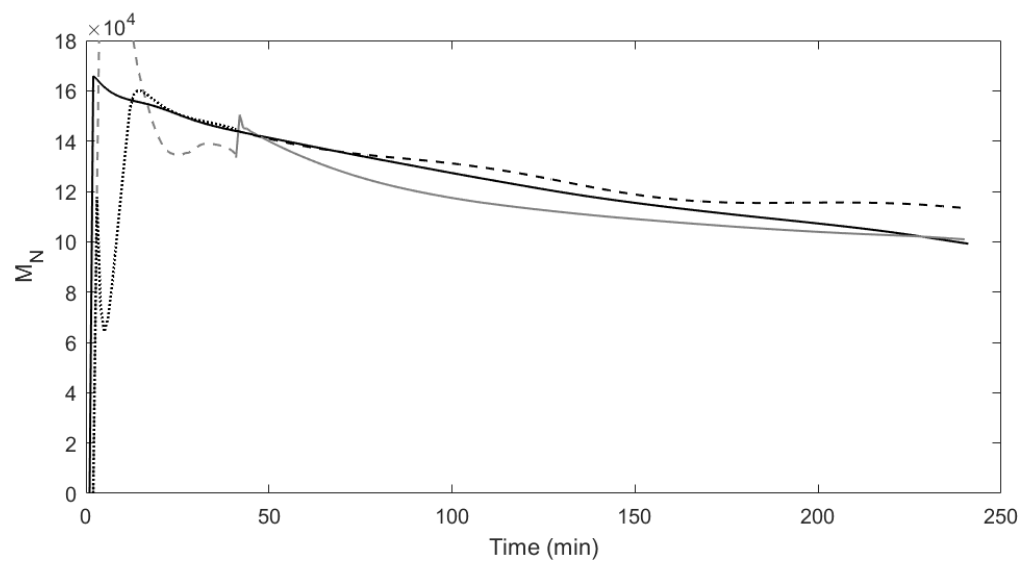


Figure 5. The number average molecular weight predictions from both models with every tenth quality measurement retained and 20% missing output data. The dashed grey lines represent estimates made by the Luenberger observer until the 40th minute and the solid grey line shows the predictions from the proposed approach in comparison to the process, represented by the solid black line. The dotted and dashed black lines show the estimates made by the Luenberger observer and traditional subspace identification using linear interpolation, respectively.

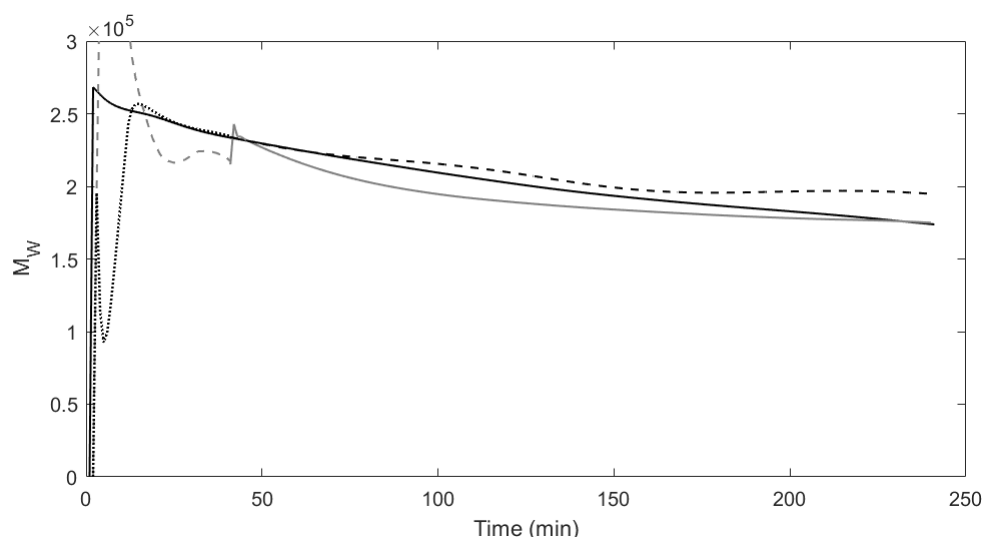


Figure 6. The weight average molecular weight predictions from both models with every tenth quality measurement retained and 20% missing output data. The dashed grey lines represent estimates made by the Luenberger observer until the 40th minute and the solid grey line shows the predictions from the proposed approach in comparison to the process, represented by the solid black line. The dotted and dashed black lines show the estimates made by the Luenberger observer and traditional subspace identification using linear interpolation, respectively.

Remark 4. Note that while the subspace identification approach itself is not a novel introduction, its application to the batch quality problem is. Quality variables are different from missing data scenarios since they are available at a much lower frequency than traditional outputs and they contain more important process information. Modeling the intermittent quality measurements along with the process outputs is a difficult task, given the large differences in data frequency. However, a singular model is desirable for control. This paper demonstrates how missing data algorithms can be adjusted to include quality variables for accurate process modeling.

4. Conclusions

In this work, a missing data subspace identification approach for batch processes with missing quality measurements and outputs is proposed. In this approach, missing quality measurements were treated as missing data, allowing the model to make predictions without inaccurately imputing the missing observations. The missing data subspace model is compared to traditional subspace identification using linear interpolation. The proposed model was able to perform better than the traditional approach for cases with quality measurements retained in intervals of 10, 20 and 30 as well as cases with random missing output data ranging from 5% to 30%. The results from the missing data approach show the benefits of using latent variable methods to handle missing data.

Author Contributions: Conceptualization, N.P. and P.M.; Data curation, N.P.; Investigation, N.P. Software, K.S.; Supervision, P.M.; Validation, K.S.; Writing—original draft, N.P.; Writing—review and editing, P.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: Financial support from the McMaster Advanced Control Consortium is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Corbett, B.; Mhaskar, P. Subspace identification for data-driven modeling and quality control of batch processes. *AIChE J.* **2016**, *62*, 1581–1601. [[CrossRef](#)]
2. Söderström, T.; Stoica, P.; Friedlander, B. An indirect prediction error method for system identification. *Automatica* **1991**, *27*, 183–188. [[CrossRef](#)]
3. Ljung, L. Prediction error estimation methods. *Circuits Syst. Signal Process.* **2002**, *21*, 11–21. [[CrossRef](#)]
4. Zhao, Y.; Huang, B.; Su, H.; Chu, J. Prediction error method for identification of LPV models. *J. Process Control* **2012**, *22*, 180–193. [[CrossRef](#)]
5. Moonen, M.; De Moor, B.; Vandenberghe, L.; Vandewalle, J. On-and off-line identification of linear state-space models. *Int. J. Control* **1989**, *49*, 219–232. [[CrossRef](#)]
6. Huang, B.; Ding, S.X.; Qin, S.J. Closed-loop subspace identification: An orthogonal projection approach. *J. Process Control* **2005**, *15*, 53–66. [[CrossRef](#)]
7. Van Overschee, P.; De Moor, B. A unifying theorem for three subspace system identification algorithms. *Automatica* **1995**, *31*, 1853–1864. [[CrossRef](#)]
8. Hu, B.; Zhao, Z.; Liang, J. Multi-loop nonlinear internal model controller design under nonlinear dynamic PLS framework using ARX-neural network model. *J. Process Control* **2012**, *22*, 207–217. [[CrossRef](#)]
9. MacGregor, J.F.; Jaeckle, C.; Kiparissides, C.; Koutoudi, M. Process monitoring and diagnosis by multiblock PLS methods. *AIChE J.* **1994**, *40*, 826–838. [[CrossRef](#)]
10. Flores-Cerrillo, J.; MacGregor, J.F. Control of batch product quality by trajectory manipulation using latent variable models. *J. Process Control* **2004**, *14*, 539–553. [[CrossRef](#)]
11. Flores-Cerrillo, J.; MacGregor, J.F. Latent variable MPC for trajectory tracking in batch processes. *J. Process Control* **2005**, *15*, 651–663. [[CrossRef](#)]
12. Qin, S.J. An overview of subspace identification. *Comput. Chem. Eng.* **2006**, *30*, 1502–1513. [[CrossRef](#)]
13. Larimore, W.E. Statistical optimality and canonical variate analysis system identification. *Signal Process.* **1996**, *52*, 131–144. [[CrossRef](#)]
14. Shang, L.; Liu, J.; Turksoy, K.; Shao, Q.M.; Cinar, A. Stable recursive canonical variate state space modeling for time-varying processes. *Control Eng. Pract.* **2015**, *36*, 113–119. [[CrossRef](#)]
15. Van Overschee, P.; De Moor, B. N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica* **1994**, *30*, 75–93. [[CrossRef](#)]
16. Verhaegen, M.; Dewilde, P. Subspace model identification part 2. Analysis of the elementary output-error state-space model identification algorithm. *Int. J. Control* **1992**, *56*, 1211–1241. [[CrossRef](#)]
17. Markovskiy, I. Exact system identification with missing data. In Proceedings of the 52nd IEEE Conference on Decision and Control, Firenze, Italy, 10–13 December 2013; pp. 151–155.
18. Ding, J.; Lin, J. Modified subspace identification for periodically non-uniformly sampled systems by using the lifting technique. *Circuits Syst. Signal Process.* **2014**, *33*, 1439–1449. [[CrossRef](#)]
19. Patel, N.; Mhaskar, P.; Corbett, B. Subspace based model identification for missing data. *AIChE J.* **2020**, *66*, e16538. [[CrossRef](#)]
20. Patel, N.; Nease, J.; Aumi, S.; Ewaschuk, C.; Luo, J.; Mhaskar, P. Integrating Data-Driven Modeling with First-Principles Knowledge. *Ind. Eng. Chem. Res.* **2020**, *59*, 5103–5113. [[CrossRef](#)]
21. Wang, J.; Qin, S.J. A new subspace identification approach based on principal component analysis. *J. Process Control* **2002**, *12*, 841–855. [[CrossRef](#)]
22. Jiang, Q.; Yan, X.; Huang, B. Performance-driven distributed PCA process monitoring based on fault-relevant variable selection and Bayesian inference. *IEEE Trans. Ind. Electron.* **2015**, *63*, 377–386. [[CrossRef](#)]
23. Ekpo, E.; Mujtaba, I.M. Evaluation of neural networks-based controllers in batch polymerisation of methyl methacrylate. *Neurocomputing* **2008**, *71*, 1401–1412. [[CrossRef](#)]
24. Fan, S.; Gretton-Watson, S.; Steinke, J.; Alpay, E. Polymerisation of methyl methacrylate in a pilot-scale tubular reactor: Modelling and experimental studies. *Chem. Eng. Sci.* **2003**, *58*, 2479–2490. [[CrossRef](#)]
25. Rho, H.J.; Huh, Y.J.; Rhee, H.K. Application of adaptive model-predictive control to a batch MMA polymerization reactor. *Chem. Eng. Sci.* **1998**, *53*, 3729–3739. [[CrossRef](#)]
26. Corbett, B.; Macdonald, B.; Mhaskar, P. Model Predictive Quality Control of Polymethyl Methacrylate. *IEEE Trans. Control. Syst. Technol.* **2013**, *23*, 3948–3953.