Control of an Industrial Scale Bioreactor using a PAT Analyser

Stephen Goldrick * **Ewan Mercer** ** **Gary Montague** *
David Lovett ** **Barry Lennox** ***

* *Biopharmaceutical Bioprocess Technology Centre, Merz Court, Newcastle
University, Newcastle-upon-Tyne. (e-mail: s.goldrick@newcastle.ac.uk)*
** *Perceptive Engineering Limited, Vanguard House, Keckwick Lane, Daresbury,
Cheshire.*
*** *Control Systems Group, School of Electrical and Electronic Engineering,
University of Manchester.*

Abstract: This work investigates the application of a "Process Analytical Technology" (PAT) analyser to control the substrate concentration over traditional sequential batch control for an industrial scale fed-batch penicillin fermentation. A simulation that utilises the historical data from four batches, where a sequential batch control strategy was implemented, was used as the benchmark reference for this comparison. The simulation accurately predicts the main outputs variables of biomass and penicillin, given the inputs from the historical data set. The simulation includes a PAT analyser, used to build a calibration model with the available off-line substrate concentration from one of the batches. The prediction from this calibration model was used as the controlled variable within a proportional integral (PI) controller to manipulate the substrate feed rate for the three remaining batches. Performance of each control strategy was analysed by comparing the final penicillin yield of each batch. An increase of 35, 20 and 9% was observed for the three batches controlled using the PI controller compared with the sequential batch control strategy.

Keywords: Industrial biotechnology; Microbial technology.

1. INTRODUCTION

Improving product yield and maintaining consistent operation is the ultimate goal in batch manufacturing. Achieving this goal in the biotechnology sector has proven to be difficult, with a major challenge being the ability to design a robust control strategy that can handle the large variations associated with batch fermentations. Major sources of variation include inconsistencies in the properties of the raw materials and the initial conditions of the batch. These variations can result in changes to important properties, such as the biomass or penicillin growth rates, which are often difficult to identify and compensate for. Further complications arise as a result of the limited on-line measurements available. Important measurements, such as substrate and product concentration are rarely available in real-time and developing a control strategy that depends on off-line measurements generally results in poor control.

Large scale batch fermentations are typically controlled by sequential, or recipe driven, batch control. This control strategy is automatic and aims to standardise the total charge volume for each process input with the overall aim of reducing variability [1]. However, this rigid control strategy can often increase variability by not responding to process disturbances or fluctuations. Also, the performance of each batch or the availability of the downstream processing unit operations is not considered using this control strategy. These factors have a major influence
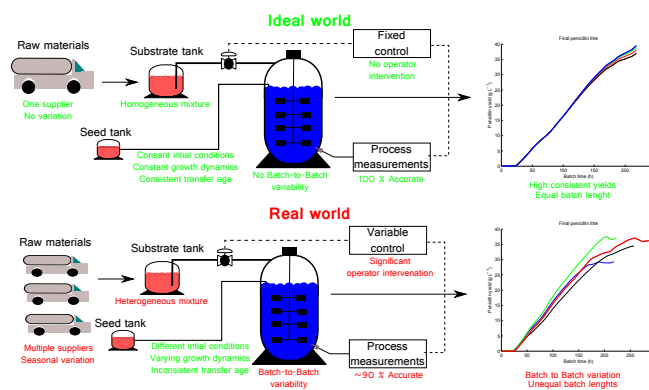

Fig. 1. Typical variation associated with batch fermentations.

on the batch length of industrial fermentations. Figure 1 summaries these practical considerations.

The current focus of this work is to analyse the benefits of process control using a PAT analyser. This technology has been shown to improve the control of batch fermentations through the timely on-line measurements of critical process parameters [2]. Similarly, [3] highlighted the use of near infra-red (NIR) spectroscopy analyser to control the substrate feed rate in a 3 litre bioreactor. This work compares the application of a simulated Raman spectroscopy sensor to control the substrate feed rate for an industrial scale bioreactor. The process control strategy using this

PAT analyser is compared, using a simulation, against the substrate feeding rate of four industrial scale fed-batch fermentations. The simulation accurately describe the industrial scale fermentations by comparing the predicted outputs against those recorded in the batch records using the main process inputs of each industrial batch.

2. MATERIALS AND METHODS

The simulation used in this work is a 1st principles mathematical model of a fed-batch 100,000 litre penicillin fermentation that includes a simulated PAT analyser in the form of a Raman spectroscopy probe. The simulation referenced as *"IndPenSim"* uses the model developed by Goldrick *et al.* ([4], [5] unpublished) which extends the penicillin fermentation models developed by Paul and Thomas [6] and Tiller *et al.* [7].

The simulation was validated using data collected from the batch records of 10 industrial scale penicillin fermentations. The results in this paper focus on data from four batches referenced as Batches 1-4. Which are a good representation of the results for the majority of the batches made available for this study. The only batches that produced inconsistent results were those that had encountered process faults, leading to batch failure. Further details of the data is available in [5]. The simulation has five inputs and six outputs and requires four batch characteristics to be initialised as summarised in figure 2.
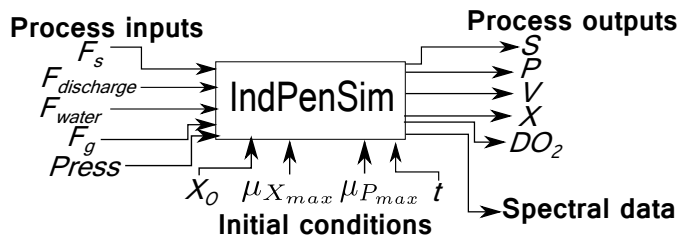

Fig. 2. Summary of the inputs and outputs for *IndPenSim*.

The inputs are the substrate feed rate (F_s), discharge rate ($F_{discharge}$), water for injection rate (F_{water}), air flow rate (F_g), head pressure ($Press$) and the outputs are the bioreactor volume (V), Raman spectroscopy data (*Spectral data*), concentrations of penicillin (P), biomass (X), substrate (S), and dissolved oxygen (DO_2). The dynamics of the simulation can be changed by adjusting the initial conditions of each batch, these include the initial biomass concentration (X_0), the maximum specific growth rate of biomass ($\mu_{X_{max}}$) and penicillin ($\mu_{P_{max}}$) and the batch length (t).

2.1 Process inputs for IndPenSim

The process inputs used in this study were taken from four industrial batch fermentations, shown in figures 3-6. The discharge rate is not displayed but was calculated from the volume profile shown in figure 7 for the historical batches. The graphs represent the typical variation associated with the process inputs for each batch. The substrate feed was the main carbon source for the fermentation and was controlled by sequential batch control that was manually

adjusted by the operator to maximise production of penicillin while ensuring the dissolved oxygen was kept above its critical value of 4 mg L^{-1}. The water for injection was used to reduce the viscosity of the broth and was manually controlled. The air flow and head pressure were controlled by the operator to ensure the dissolved oxygen concentration was kept above its critical value. The discharge rate ensured that the volume of the bioreactor was kept below its set-point of 100,000 litres. Further details of these control variables can be found in [5].
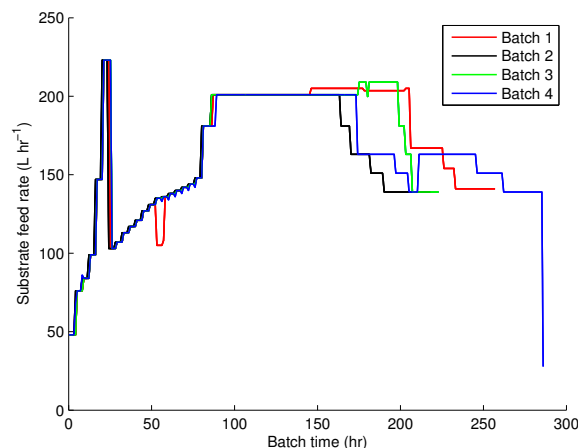

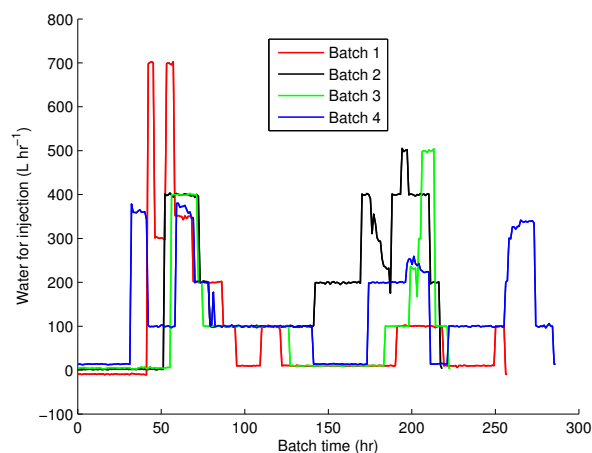Fig. 3. Substrate feeding rates for four industrial batches.

Fig. 4. Water for injection rates of four industrial batches.

2.2 Initial conditions

The initial conditions for the four industrial batches are displayed in table 1. The initial biomass concentration

Batch reference	X_0 (g L-1)	$\mu_{X_{max}}$ (hr^{-1})	$\mu_{P_{max}}$ (hr^{-1})	Batch length (hr)
Batch 1	1.65	0.25	0.0409	257
Batch 2	2.6836	0.2349	0.0439	218
Batch 3	3.1105	0.2366	0.0459	223
Batch 4	1.74	0.3124	0.0434	286

Table 1. Data calculated for four industrial batches and used as the inputs for *IndPenSim*.
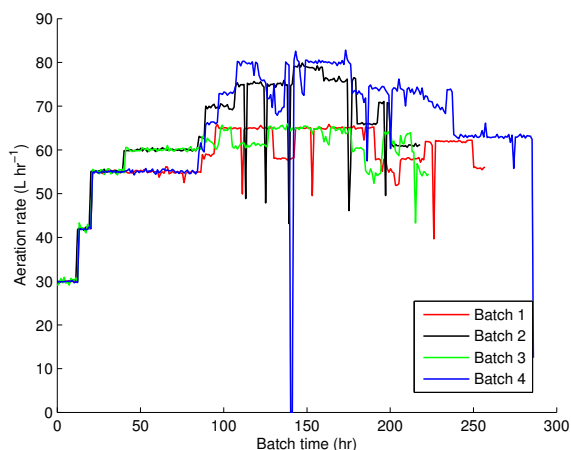
and the maximum specific growth rate of biomass were
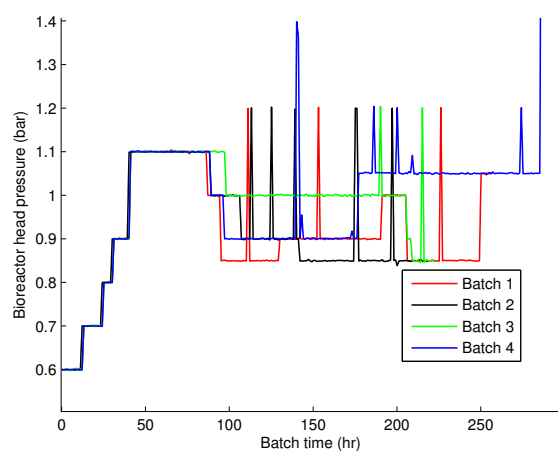
Fig. 5. Air flow rate for four industrial batches.

Fig. 6. Vessel pressure for four industrial batches.

Fig. 7. Recorded volume for four industrial batches compared to the simulation volume using *IndPenSim* (·−historical data,−− simulated data).

Fig. 8. Comparison of the simulated (s) penicillin concentration, shown by the solid lines, against the historical penicillin concentration, shown by the dashed lines, of the four batches (·−historical data,−− simulated data).

calculated using the oxygen uptake rate from each industrial batch. The maximum specific growth rate of penicillin was calculated using the first three off-line measurements of penicillin which were recorded every 24 hours. The batch length was the total time of the fermentation. The initial conditions and batch length was kept constant in the comparison of the two different control strategies.

2.3 Process outputs

IndPenSim uses the process inputs and batch characteristics to predict the substrate, penicillin, biomass, volume and DO_2 for each batch. The recorded off-line penicillin concentration and volume of the four industrial batches is compared against the simulated values predicted using *IndPenSim* as shown in figures 7 and 8. The substrate prediction is discussed in section 3 and the prediction of biomass and dissolved oxygen is outside of the scope of this paper.

2.4 Spectral data

IndPenSim contains a simulation of a Raman spectroscopy sensor. The spectra are generated by applying six Gaussian distributions, three of which are related to the concentration of penicillin and the remaining three are related to the substrate concentration. These Gaussian distributions
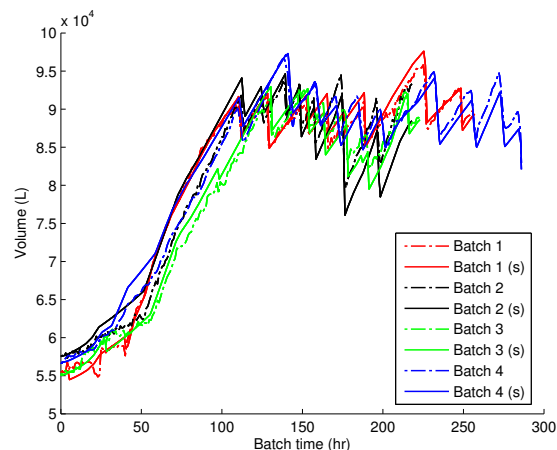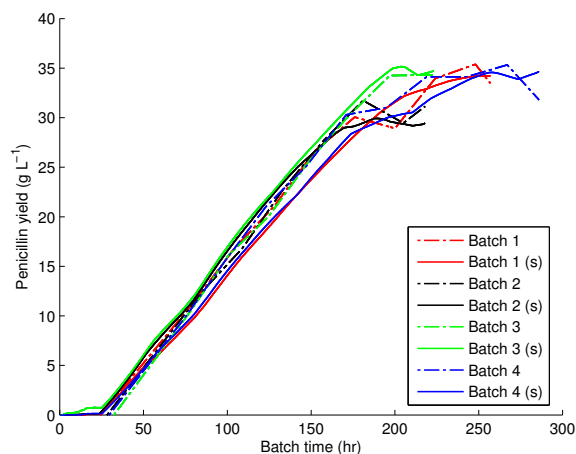
were then added to a baseline spectra. The intensity of baseline spectra increased as a function of the biomass concentration, to simulate fluorescence. A linear relationship was assumed between peak high and substrate or penicillin concentration as defined by the Beer-Lamber law. The spectra was recorded every 12 minutes and can be used to model the substrate concentration on-line provided a calibration model is successfully built. An example of the raw spectra collected for Batch 1 is shown in figure 9.

2.5 Development of a calibration model

The simulated substrate concentration combined with the spectral data from Batch 1 were used to build a calibration model to predict the substrate concentration. It was assumed that only the off-line measurements of the substrate concentration, measured every four hours, were available for the calibration model. The model was built using the off-line substrate data (S_{OF}), consisting of a vector with 85 rows, and the corresponding spectral data (X_{Spec}) consisting of a matrix of 2200 columns and 85 rows. The
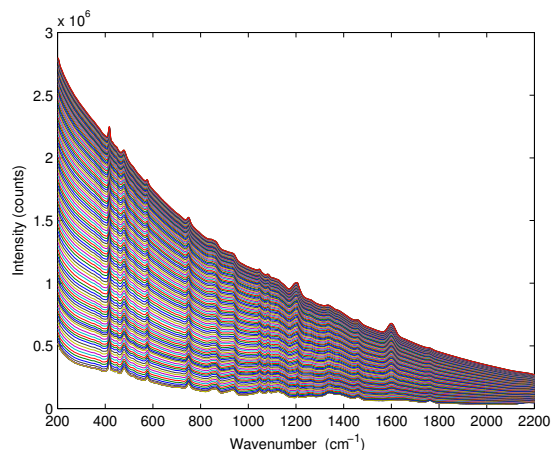
Fig. 9. Spectra collected from Batch 1
.

Fig. 10. Regression coefficients $\hat{B}$, calculated for the first three latent variables of the PLS model for Batch 1.

column vectors of the matrix represent the wavenumbers of the spectra. The wavenumbers of interest that contain information related to the substrate concentration were chosen using the regression coefficients (B) of a PLS model taking X_{Spec} as the X-block and S_{OF} as the Y block. The PLS model was built using the technique outlined in [8]. The spectral data was first decomposed in to its principal components, generating a matrix of scores, T, and loadings, P. The off-line substrate concentration was decomposed in a similar fashion generating a matrix of scores, U, and loadings, Q, as shown below:

$$X_{Spec} = TP' + E \qquad (1)$$

$$S_{OF} = UQ' + F \qquad (2)$$

The inner relationship that relates the scores of the X_{Spec} block to the scores of the S_{OF} block was calculated as follows:

$$U = TW \qquad (3)$$

This inner relationship of regression coefficients was improved by exchanging the scores, T and U, in an iterative calculation. This allows information from one block to be used to adjust the orientation of the latent vectors in the other block and vice versa. Once the complete model was calculated, a matrix of regression vectors, $\hat{B}$, was generated:

$$\hat{B} = P(P'P)^{-1}WQ' \qquad (4)$$

$$\hat{S} = X_{Spec}\hat{B} \qquad (5)$$

Where $\hat{S}$ is the predicted substrate concentration. The magnitude of regression vectors ($\hat{B}$) was used to determine the wavenumbers of the spectra that contain information about the substrate concentration. The wavenumbers with large values of $\hat{B}$ show that there is a relation between those wavenumbers (X_{Spec}) and the substrate concentrations (S_{OF}). These wavenumbers can be selected by plotting the first three latent variables of $\hat{B}$ as shown in figure 10. The wavenumbers of interest were chosen from this figure and were taken as: 170:270 620:660 953:1153 cm^{-1}. The off-line substrate concentrations was divided into a calibration and validation data set as shown in figure 11. The growth phase of the fermentation considered as the first 30 hours of each batch was ignored due to the high substrate concentration during this period. The spectral data was pre-processed using a standard Savitzky-
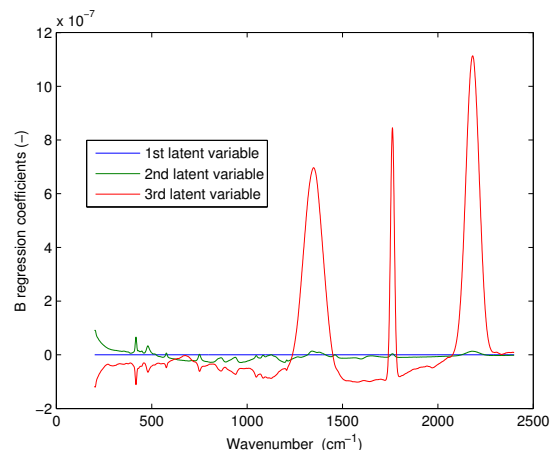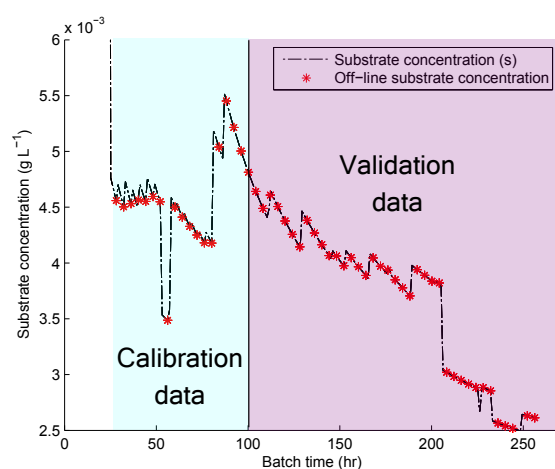

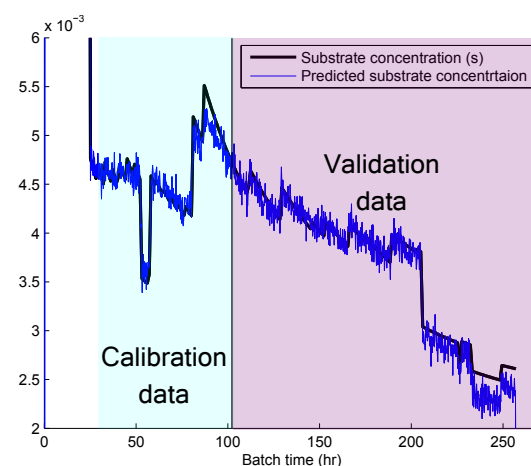Fig. 11. Calibration and validation data split for batch 1.

Fig. 12. Predicted substrate concentration compared the actual simulated substrate concentration for Batch 1.

Golay smoothing technique using a 15 point average and taking the first derivative. A partial least squares model was then developed using the PLS technique described earlier. Five latent variables were chosen for the model as it accurately described the substrate concentration for both the calibration and validation data sets of Batch 1 as shown in figure 12.

2.6 PAT Controller

The calibration model was applied on-line and utilised for the prediction of the substrate concentration for Batches 2-4. This predicted substrate concentration was first filtered using a three point moving average filter to reduce the noise associated with the prediction. This filtered signal was used as the controlled variable in a proportional integral (PI) controller that manipulated the substrate feed rate. The PI controller was switched on after 100 hours which was taken as the transition point from the growth phase to the stationary phase for each batch. The set-point of the controlled variable was 5.5×10^{-3} g L^{-1}. This set-point was taken as the average of the off-line substrate concentrations as shown in figure 11. The volume of the tank was controlled by manipulating the discharge flow rate to ensure that the volume did not exceed 90,000 litres. The controller set-up is illustrated in figure 13.
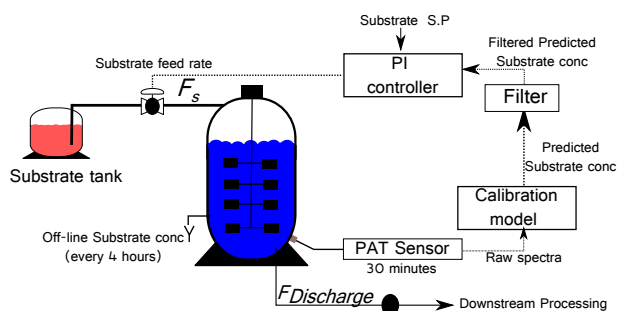

Fig. 13. Overview of PAT controller

3. DISCUSSION

The mathematical simulation, *IndPenSim*, was used to test the performance of a control strategy using a PAT analyser on four industrial scale penicillin fermentations with varying dynamics. Before the proposed PAT controller was applied, the accuracy of the simulation was first validated. Validation was made by using the process inputs and initial batch conditions from the four historical batches as model inputs in the simulation and then comparing the predicted outputs against the actual responses. Figure 8 compares the actual penicillin measurements, recorded using off-line laboratory analysis, against the simulated values of penicillin by *IndPenSim*. This figure shows that the simulation tracks the off-line measurements very closely for each of the four batches. The simulation was also shown to predict the volume in the bioreactor accurately based on the inputs and the rate of discharge as highlighted in figure 7.

Having validated the simulation, the performance of the proposed PI controller system was assessed. Before the control strategy could be implemented a calibration model was built using the data from Batch 1 as discussed in section 2.5. This calibration model gave a good prediction of the substrate concentration as shown in figure 12. Although the prediction is shown to be noisy the model accurately predicts the changes in the substrate concentration throughout the entire batch. In order to reduce the effect of this noise on the PI controller the input signal was filtered using the technique discussed
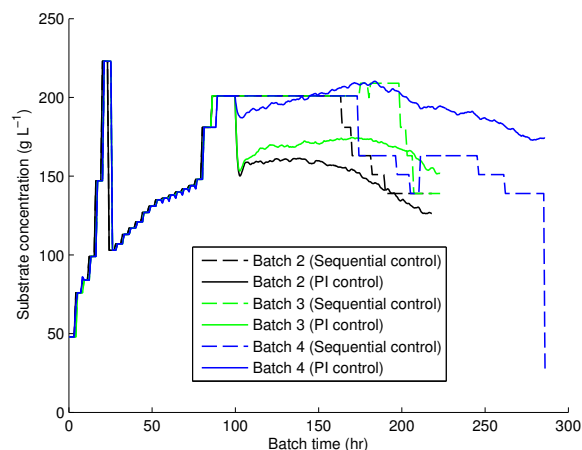

Fig. 14. Comparison of the substrate feed rate using the APC solution against how it was controlled using through operator control
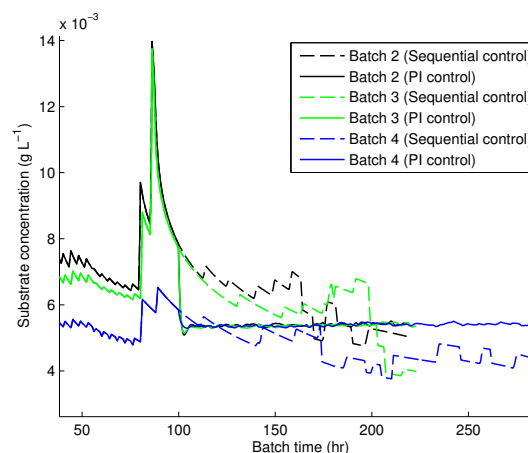

Fig. 15. Comparison of the substrate concentration when the substrate feed rate is controlled using a PI controller against and when controlled by an operator.

in section 2.6. The substrate feed rate manipulated by the PI controller is compared against the previous operator controlled substrate recorded in the batch records for Batches 2-4 as shown in figure 14. The controller is switched on after 100 hours and reduces the substrate feed rate to control the substrate concentration to its set-point of 5×10^{-1} g L^{-3}. This set-point is tightly controlled for the remainder of each of the batches. This tight control on the substrate concentration is shown to improve the production of penicillin for each batch as seen in figure 17. The final penicillin yield was shown to increase by 35%, 20 % and 9% for Batches 2, 3 and 4, respectively. The volume control for each batch using the two different control strategies is highlighted in figure 17. The volume control using the PI controller strategy is shown to control the level of the bioreactor very tightly compared to when it was controlled using sequential batch control. Tight control is highly desirable in large scale batch fermentations as the large changes observed during operator controlled volume can represent disturbances on other process variables and reduce product concentration.
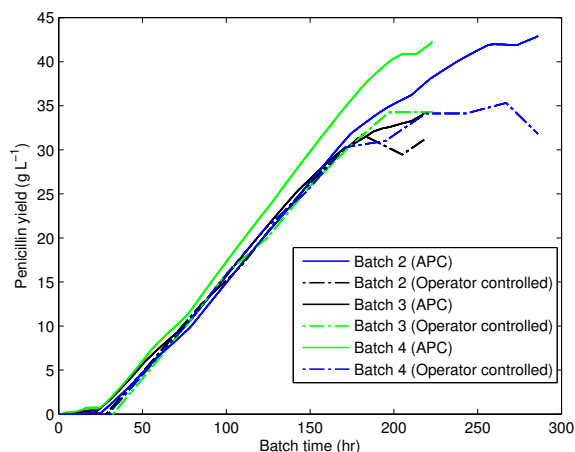
Fig. 16. Comparison of penicillin concentrations using the APC controller compared to when it is controller by the operator.
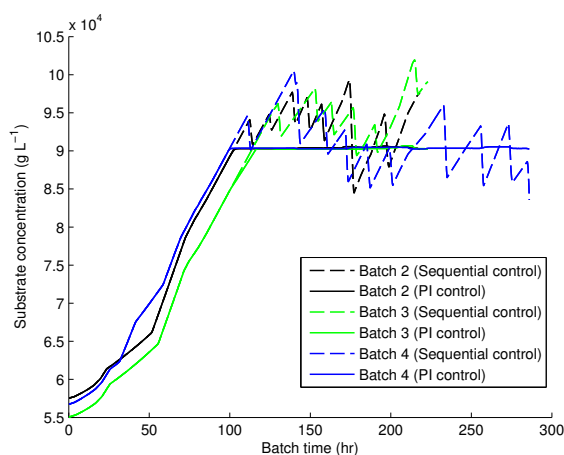

Fig. 17. Penicillin concentration 1.

4. CONCLUSIONS

The benefits of using a PAT analyser to help control the substrate feed rate for an industrial scale penicillin fermentation are highlighted here. Through the application of this advanced technology, the substrate was measurement online. These measurements were used by a PI controller to manipulate the substrate feed rate. By controlling the substrate feed concentration to a fixed set-point, the penicillin production rate was shown to increase compared to when the batch was controlled using sequential batch control.

ACKNOWLEDGEMENTS

REFERENCES

[1] Boudreau A. and McMillan K., Gregory, *New directions in bioprocess modelling an control: Maximising process analytical technolology benifits.* ISA, 2007

[2] E.K. Read, J.T. Park, R.B. Shah, B.S. Riley, K.A. Brorson, A.S. Rathore *Process Analytical Technology (PAT) for Biopharmaceutical Products: Part I.* *Concepts and Applications* Biotechnology and Bioengineering, Vol. 105, No. 2, February 1, 2010

[3] Marian Navrátila, Anna Norbergb, Lars Lembrénb, Carl-Fredrik Mandeniusa *On-line multi-analyzer monitoring of biomass, glucose and acetate for growth rate control of a Vibrio cholerae fed-batch cultivation* Journal of Biotechnology 115 (2005) 67-79

[4] Stephen Goldrick, Barry Lennox, David Lovett, Keith Smith, Gary Montague *The Development of a Simulation to Address the Real Challenges Associated with Industrial Scale Penicillin Production.* Computer Applications in Biotechnology conference, Mumbai Dec 2013

[5] Stephen Goldrick, Andrei Stefan, Gary Montague, David Lovett, and Barry Lennox *The development of a simulation to address the pratical challenges associated with industrial scale penicillin production* Biotechnology and Bioengineering, manuscript in preparation to be submitted in Jan 2014

[6] G. C. Paul and C. R. Thomas, *A structured model for hypal differentiation and Penicillin Production Using Penicillium chrysogenum* Biotechnology and Bioengineerng, 51,558-572 1996.

[7] Volker Tiller, Juliane Meyerhoff, Ditmar Sziele, Karl Schgerl, Karl-Heinz Bellgardt , *Segregated mathematical model for the fed-batch cultivation of a high-producing strain of Penicillium chrysogenum* Journal of Biotechnology, 34, 2, 15, 119-131, 1994

[8] Paul Gemperline *Pratical Guide to Chemometrics* Taylor & Francis Group, 2006.