MODEL BASED APPROACHES IN CONTINUOUS DOWNSTREAM PURIFICATION OF BIOTHERAPEUTICS

ANAMIKA TIWARI

DEPARTMENT OF CHEMICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

JULY 2024

MODEL BASED APPROACHES IN CONTINUOUS DOWNSTREAM PURIFICATION OF BIOTHERAPEUTICS

by

ANAMIKA TIWARI

DEPARTMENT OF CHEMICAL ENGINEERING

Submitted

in fulfilment of the requirements of the degree of Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

JULY 2024

Dedicated to my family

Certificate

This is to certify that the thesis entitled **"MODEL BASED APPROACHES IN CONTINUOUS DOWNSTREAM PURIFICATION OF BIOTHERAPEUTICS"** being submitted by **ANAMIKA TIWARI** to the Indian Institute of Technology Delhi for the award of the degree of **Doctor of Philosophy** is a record of the original bonafide research work carried out by her under my guidance and supervision. The results contained in this thesis have not been submitted in part or in full to any other University or Institute for the award of any degree or diploma.

I certify that she has pursued the prescribed course of research.

Prof. Anurag S. Rathore
Department of Chemical Engineering
Indian Institute of Technology Delhi

Acknowledgements

This thesis would not have been feasible without the invaluable assistance of numerous persons who contributed in many ways. I would like to express my profound gratitude and sincere appreciation to them.

First and foremost, I want to express my heartfelt gratitude to my Professor, Dr. Anurag S. Rathore for his insightful discussions, invaluable advice, unwavering support, and patience during my doctoral program. His treasured support was influential in helping me overcome numerous obstacles and complete my research work. He has helped me not only as a researcher, but also emotionally, giving me plenty of moral support and freedom.

I would like to extend my sincere thanks to my PhD research committee members, Prof. James Gomes, Prof. Anupam Shukla and Prof. Manojkumar Ramteke whose periodic assessment and review of my work helped in going forward in the right direction. Their feedback was quite helpful in building a conceptual strategy for approaching research problems in a systematic manner.

I would also like to express profound appreciation to the Tata Consultancy Services Research Scholar Program for providing me with a research scholarship as well as technical support for my doctorate studies.

I am extremely grateful to my colleagues and seniors in the lab, Dr. Lalita Kanwar Shekhawat Dr. Nikhil Kateja, Dr. Vishwanath Hebbi, Dr. Garima Thakur, Vikrant Bansode, Anupa Pahuja, Nitika, Shantanu Banerjee, Himanshu Malani and Rucha Patil for their helpful discussions, encouragement and cooperation for smooth completion of my work. I would also like to thank all the other members of the bioprocessing lab for their kind assistance and encouragement on regular basis.

Special thanks to my dear friend and roommate Dr. Akanksha Mishra for her selfless love, care and assistance in every way possible.

No words can explain how grateful I am to all my family members including my parents, sister, brother-in-law, husband, and in-laws for consistently praying and motivating me to achieve success at every stage of my life. Special thanks to my niece for always putting a smile on my face and being a great source of energy. I am indebted to all of them for their blessings, love, care, and faith for successful completion of this work.

ANAMIKA TIWARI

Abstract

The biopharma industry traditionally uses the batch model for manufacturing biotherapeutics due to risk tolerance and regulatory concerns. However, there is a growing interest in continuous processing to meet rising demand. Continuous processing offers benefits such as improved productivity, consistent product quality, and reduced costs. While the majority of research focus has been on the creation of end-to-end continuous biomanufacturing processes in recent years, limited attention has been paid to real time monitoring and control of continuous processes. The aim of this thesis is to develop model based designs and control of downstream purification steps along with development of appropriate online Process Analytical Technology (PAT) tool for real time monitoring of Critical Quality Attributes (CQAs) in continuous processing of biotherapeutics.

The first objective of the work was to explore applications of machine learning (ML) algorithms for qualitative monitoring and quantitative prediction of column state and performance for resin cycling study of continuous capture chromatography. This process is crucial in monoclonal antibody purification due to resin costs. Advanced ML algorithms were proposed to detect column integrity breaches and predict yield decline in protein A chromatography. The developed algorithms successfully predicted anomalies by detecting subtle changes in column packing and decline in yield, enabling real-time control decisions for preventive measures.

The second objective of the work was to model the complex resin decay phenomena due to extended re-use of protein A resin in continuous capture chromatography using general rate model (GRM) for monoclonal antibody (mAb). Our findings show that pore and film diffusion slow down with each cycle, affecting the protein transfer rate. It was also observed that although pore diffusion was the limiting step, film diffusion was also slowing with number of cycles leading to slower diffusion rate of protein out of solid phase to bulk phase. The methodology and findings proposed in this work can assist in predicting reliable performance of Protein A resin with useful insights regarding the underlying mechanism behind resin fouling.

Besides process modeling, creation of appropriate PAT tool is required for real time monitoring of product quality attributes in continuous processing, thereby ensuring product quality and consistency. In this objective, a cost-effective approach for using a traditional offline high-performance liquid chromatography (HPLC) for online analysis using a 2 way/6 port valve to facilitate simultaneous automated sampling of product stream eluting from a process column and fractionation was proposed. The versatility of the proposed online configuration has been verified by demonstrating its use for separation of aggregates and separation of charge variants

of mAb. Furthermore, predictive modeling was implemented to facilitate pooling to achieve the desired purity. The study demonstrates that the proposed online HPLC configuration could be used for PAT applications in continuous preparative chromatography.

Finally, model based design and control of end-to-end integrated downstream purification process coupled with online HPLC PAT tool for achieving robust pooling decision making in continuous processing of mAb was proposed. Robust pooling decision is required for achieving uniform charge variant composition for mAbs. In this study, we proposed an end-to-end modeling of continuous downstream process providing for digital twin, integrated with an online HPLC-PAT tool for delivering real time pooling decisions to achieve uniform charge variant composition. The model states were updated in real time using online HPLC charge variant data for prediction of the initial and final cut point for cation exchange chromatography (CEX) eluate, according to which the process chromatography was directed to switch from collection to waste to achieve the desired charge variant composition in the CEX pool. The control strategy was successfully implemented for wide range of charge variants in the feed with its composition in the range of 13-17% for acidic, 18-23% for main and 59-68% for basic variants. All the case studies yielded CEX pool of uniform distribution of charge variants with process yield for main species was >85%. The results thus successfully demonstrate the technical feasibility of creating digital twins of bioprocess operations and their utility for process control.

In order to run a reliable, long-term continuous process with increased efficiency and robustness, we believe the studies presented here as a part of this thesis will promote development of several modeling based techniques together with application of appropriate PAT tools for real time monitoring and control of bioprocess unit operations.

सार

जोखिम सहयता और विनियामकीय कारणों से बायोफार्मा उद्योग परंपरागत रूप से बायोथेरेप्यूटिक्स के निर्माण के लिए बैच मॉडल का उपयोग करता है। हालाँकि, बढ़ती माँग को पूरा करने के लिए अविरत प्रसंस्करण में रुचि बढ़ रही है। अविरत प्रसंस्करण से बेहतर उत्पादकता, सुसंगत उत्पाद गुणवत्ता और कम लागत जैसे लाभ मिलते हैं। जहां पर हाल के वर्षों में अनुसंधान का अधिकांश ध्यान एंड-टू-एंड अविरत जैव-विनिर्माण प्रक्रियाओं के निर्माण पर रहा है, तात्कालिक निगरानी और अविरत प्रक्रियाओं के नियंत्रण पर सीमित ध्यान दिया गया है। इस थीसिस का उद्देश्य बायोथेरेप्यूटिक्स के अविरत प्रसंस्करण में क्रिटिकल क्वालिटी ऐट्रिब्यूट्स (सीक्यूए) की तात्कालिक निगरानी के लिए उपयुक्त ऑनलाइन प्रोसेस एनालिटिकल टेक्नोलॉजी (पीएटी) उपकरण के विकास के साथ-साथ डाउनस्ट्रीम शुद्धिकरण चरणों के मॉडल आधारित डिजाइन का विकास और नियंत्रण करना है।

कार्य का पहला उद्देश्य कॉलम की स्थिति व प्रदर्शन की गुणात्मक निगरानी और मात्रात्मक अनुमानन और अविरत कैप्चर क्रोमैटोग्राफी के रेज़िन साइक्लिंग अध्ययन के लिए मशीन लर्निंग (एमएल) एल्गोरिदम के अनुप्रयोगों का पता लगाना था। रेज़िन लागत के कारण मोनोक्लोनल एंटीबॉडी शुद्धिकरण में यह प्रक्रिया महत्वपूर्ण है। कॉलम अखंडता विच्छेदों का पता लगाने और प्रोटीन 'A' क्रोमैटोग्राफी में यील्ड में गिरावट की भविष्यवाणी करने के लिए उन्नत एमएल एल्गोरिदम प्रस्तावित किए गए हैं। विकसित एल्गोरिदम द्वारा कॉलम पैकिंग और यील्ड में गिरावट में सूक्ष्म परिवर्तनों का पता लगाकर विसंगतियों की सफलतापूर्वक भविष्यवाणी की गई, जिससे निवारक उपायों के लिए तात्कालिक नियंत्रण निर्णय लेने में मदद मिली।

कार्य का दूसरा उद्देश्य, मोनोक्लोनल एंटीबॉडी (एमएबी) के लिए जनरल रेट मॉडल (जीआरएम) का उपयोग करके अविरत कैप्चर क्रोमैटोग्राफी में प्रोटीन 'A' रेज़िन के विस्तारित पुनरुपयोग के कारण होने वाली रेज़िन क्षयन की जटिल परिघटना को मॉडल करना था। हमारे निष्कर्ष बताते हैं कि प्रत्येक चक्र के साथ पोर और फिल्म विसरण धीमा हो जाता है, जिससे प्रोटीन स्थानांतरण दर प्रभावित होती है। यह भी देखा गया कि यद्यपि पोर विसरण दर-नियंत्रक चरण था, फिल्म विसरण भी चक्रों की संख्या के साथ धीमा हो रहा था, जिससे ठोस अवस्था से सतत द्रव अवस्था तक प्रोटीन की विसरण दर धीमी होती गई थी। इस कार्य में प्रस्तावित पद्धति और निष्कर्ष, रेज़िन के दूषित होने के पीछे अंतर्निहित क्रियाविधि के बारे में उपयोगी जानकारी के साथ प्रोटीन 'A' रेज़िन के प्रदर्शन के बारे में विश्वसनीय अनुमान करने में सहायता कर सकते हैं।

प्रक्रिया मॉडलिंग के अलावा, अविरत प्रसंस्करण में उत्पाद की गुणवत्ता विशेषताओं की तात्कालिक निगरानी के लिए उपयुक्त पीएटी उपकरण का निर्माण आवश्यक है, जिससे उत्पाद की गुणवत्ता और स्थिरता सुनिश्चित हो सके। इस उद्देश्य में, एक लागत-कुशल दृष्टिकोण प्रस्तावित किया गया है जिसमें एक पारंपरिक ऑफलाइन हाई-परफॉरमेंस लिक्विड क्रोमैटोग्राफी (एचपीएलसी) का उपयोग ऑनलाइन विश्लेषण के लिए 2-तरफा/6-पोर्ट वाल्व के साथ किया जाता है, ताकि प्रक्रिया कॉलम से उत्पाद स्ट्रीम की स्वचालित सैंपलिंग और फ्रैक्शनेशन की सुविधा एक साथ मिल सके। प्रस्तावित ऑनलाइन कॉन्फ़िगरेशन की बहुपरकारीता को एग्रीगेट्स और एमएबी के चार्ज वेरिएंट्स के पृथक्करण में इसके उपयोग से सत्यापित किया गया है। इसके अलावा, वांछित शुद्धता प्राप्त करने के लिए पूलिंग को सुविधाजनक बनाने हेतु प्रीडिक्टिव मॉडलिंग का प्रयोग किया गया। अध्ययन यह दर्शाता है कि प्रस्तावित ऑनलाइन एचपीएलसी कॉन्फ़िगरेशन को अविरत प्रेपरेटिव क्रोमैटोग्राफी में पीएटी अनुप्रयोगों के लिए उपयोग किया जा सकता है।

अंत में, एंड-टू-एंड एकीकृत डाउनस्ट्रीम शुद्धिकरण प्रक्रिया का मॉडल आधारित डिज़ाइन और नियंत्रण प्रस्तावित किया गया, जो अविरत एमएबी प्रोसेसिंग में ठोस पूलिंग निर्णय लेने के लिए ऑनलाइन एचपीएलसी- पीएटी उपकरण के साथ जोड़ा गया। एमएबी के लिए एकसमान चार्ज वेरिएंट संघटन प्राप्त करने के लिए ठोस पूलिंग निर्णय आवश्यक है। इस अध्ययन में, हमने एक डाउनस्ट्रीम प्रक्रिया के एंड-टू-एंड मॉडलिंग का प्रस्ताव किया, जो डिजिटल ट्विन के लिए प्रस्तुत किया गया, जो ऑनलाइन एचपीएलसी- पीएटी उपकरण के साथ एकीकृत है, जो तात्कालिक पूलिंग निर्णय प्रदान करने के लिए है ताकि एकसमान चार्ज वेरिएंट संघटन हासिल की जा सके। कैटायन एक्सचेंज क्रोमैटोग्राफी (सीईएक्स) एलुएट के लिए प्रारंभिक और अंतिम कट बिंदु के पूर्वानुमान के लिए ऑनलाइन एचपीएलसी चार्ज वेरिएंट डेटा का उपयोग करके मॉडल दशाओं को तात्कालिक अपडेट किया गया, जिसके अनुसार वांछित चार्ज वेरिएंट संघटन को प्राप्त करने के लिए प्रक्रिया क्रोमैटोग्राफी को संग्रह से अपशिष्ट मोड में स्विच करने के लिए निर्देशित किया गया। नियंत्रण रणनीति को फीड में चार्ज वेरिएंट्स की विस्तृत रेंज के लिए सफलतापूर्वक लागू किया गया जिसमें एसिडिक की रेंज 13-17%, मेन की रेंज 18-23% और बेसिक वेरिएंट की रेंज 59-68% थी। सभी घटना अध्ययनों ने साबित किया कि सीईएक्स पूल में चार्ज वेरिएंट्स का एकसमान वितरण हुआ और मेन वेरिएंट के लिए प्रक्रिया उत्पादन >85% था। इस प्रकार ये परिणाम, जैवप्रक्रिया संचालन के डिजिटल ट्विन्स बनाने और उनके प्रक्रिया नियंत्रण के लिए उपयोगी होने की तकनीकी संभावना को सफलतापूर्वक प्रदर्शित करते हैं।

एक विश्वसनीय, दीर्घकालिक अविरत प्रक्रिया को अधिक दक्षता और मजबूती के साथ चलाने के लिए, हम विश्वास करते हैं कि इस थीसिस में प्रस्तुत अध्ययन कई मॉडलिंग आधारित तकनीकों के विकास के साथ-साथ जैव-प्रक्रिया इकाई संचालनों के तात्कालिक मॉनिटरिंग और नियंत्रण के लिए उचित पीएटी उपकरणों के अनुप्रयोग को बढ़ावा देगा।

Contents

Certificate ..... **i**

Acknowledgements ..... **ii**

Abstract ..... **iii**

List of Figures ..... **x**

List of Tables ..... **xiii**

List of Abbreviations ..... **xiv**

1 Introduction **1**

1.1 Background 2

1.2 Problem definition, scope and objectives of research 4

1.2.1 Machine learning based algorithms for qualitative monitoring and quantitative prediction of column state and performance in continuous capture chromatography 4

1.2.2 Mechanistic modelling of protein A resin fouling in continuous capture chromatography 4

1.2.3 Creation of online PAT tool in conjunction with predictive modeling as an enabler for continuous processing 5

1.2.4 Model based design and control of end-to-end integrated downstream process coupled with online HPLC PAT tool in continuous processing of mAbs 5

2 Literature survey **6**

2.1 Continuous processing for manufacturing of biotherapeutics 7

2.2 Model based approaches for continuous capture chromatography 8

2.3 Model based approaches for continuous polishing chromatography 13

2.4 Model based approaches for integrated process 14

2.5 Process analytical technology for continuous processing 17

2.6 Summary of literature survey 19

3 Machine learning based algorithms for qualitative monitoring and quantitative prediction of column state and performance in continuous capture chromatography 21

3.1 Introduction 22

3.2 Materials and methods 25

3.2.1 Materials and equipment 25

3.2.2 Protein A PCC chromatography method 25

3.3 Machine learning algorithms for detecting anomaly in column packing quality 27
3.3.1 Pre-processing of chromatography data 27
3.3.2 Modeling techniques using PCA-IF and LSTM AE 27
3.4 Machine learning algorithms for yield detection 32
3.4.1 PLS model 32
3.4.2 ANN and PLS-ANN augmented model 32
3.5 Results and Discussion 33
3.5.1 Anomaly detection using PCA-IF and LSTM AE 33
3.5.2 PLS, ANN and augmented PLS-ANN model for yield prediction 39
3.6 Conclusions 42
4 Mechanistic modeling of protein A resin fouling in continuous capture chromatography 44
4.1 Introduction 45
4.2 Materials and methods 47
4.2.1 Materials and equipment 47
4.2.2 Protein A PCC chromatography method 47
4.3 Mechanistic model 48
4.3.1 Protein A chromatography column model 48
4.3.2 Estimation of model parameters 51
4.4 Results and Discussion 52
4.4.1 Simulation results 53
4.4.2 Analysis of concentration profiles inside the column 57
4.4.3 Analysis of driving forces for mass transfer 58
4.5 Conclusions 60
5 Creation of online PAT tool in conjunction with predictive modeling as an enabler for continuous processing 62
5.1 Introduction 63
5.2 Materials and Methods 64
5.2.1 Materials 64
5.2.2 Case Study I: Analysis of aggregates 64
5.2.3 Case Study II: Analysis of charge variants 65
5.3 Online HPLC 65
5.4 Systematic Interpolation Technique of online HPLC data for pooling decision making 66

5.5 Results and Discussion 68
5.5.1 Case Study I: Separation of mAb aggregates by cation exchange (CEX) process chromatography 68
5.5.2 Case study II: Separation of charge variants by cation exchange (CEX) process chromatography 71
5.5.3 Integration of online HPLC with multicolumn chromatography 73
5.5.4 Predictive modeling of online HPLC data for PAT application 74
5.5.5 Benefits of proposed configuration 76
5.6 Conclusions 76
6 Model based design and control of end-to-end integrated downstream process coupled with online HPLC PAT tool in continuous processing of mAbs 78
6.1 Introduction 79
6.2 Materials and methods 81
6.2.1 Materials 81
6.2.2 Methods 82
6.2.3 Process modeling 88
6.3 Results and discussion 94
6.3.1 Parameter estimates and model predictions for Protein A PCC 95
6.3.2 Parameter estimates and mechanistic model predictions of CEX 97
6.3.3 Empirical vs. mechanistic modeling 100
6.3.4 Control strategies for adaptive pooling 100
6.3.5 Integrated continuous operation for steady state and perturbation runs 102
6.4 Conclusions 106
7 Conclusions and Scope of Future Work 109
7.1 Conclusions 110
7.2 Scope of future work 111
References 114
Bio-Data 129

List of Figures

Figure 1.1 Conceptual framework of digital twin for biomanufacturing of monoclonal antibodies 3

Figure 3.1 Conceptual framework of (A) hybrid model (PCA followed by IF) and (B) LSTM autoencoder for detecting column integrity breach. 30

Figure 3.2 An illustration of the LSTM Autoencoder network. 31

Figure 3.3 Scree plot of PCA analysis explaining % variance for each principal component. 34

Figure 3.4 IF score plots for (A) training dataset:1-24 cycles, (B) testing dataset: 25-40 cycles, and (C) validation dataset: 41-76 cycles. 35

Figure 3.5 (A) Convergence of loss function of LSTM model, (B) and (C) True and predicted values of scaled UV absorbance values for 23rd and 24th cycle of training dataset. 37

Figure 3.6 Mean block error plots (A) trained dataset (1-24 cycles), (B) testing dataset (25-40 cycles) and (C) validation dataset (41-76 cycles). 38

Figure 3.7 Prediction error vs. experimental yield plot using (A) PLS model (B) ANN and (C) PLS-ANN model. 41

Figure 4.1 Methodology for modelling of fouling phenomena for protein A chromatography. 52

Figure 4.2 Comparison of experimental (solid lines) and simulated elution profiles (dashed lines) for 0th to 50th cycles (A-F) along with pH variation. 54

Figure 4.3 Extrapolated plot (solid line) of variation of mass transfer parameters using estimated (scattered) values of ξ, π (A-B) and (C) experimental εp. 55

Figure 4.4 Comparison of experimental (solid lines) and predicted elution profiles (dashed lines) for 50th to 80th cycles (A-C) along with pH variation. 56

Figure 4.5 (A) Solid phase protein concentration profiles for cycles for 0^{th} to 80^{th} cycles at mid length ($Z = L/2$) of the column, (B) Normalized differential of solid phase protein concentration along with zoomed view (from 100-200s) of the plot highlighting the variation in skewness for 0^{th} to 80^{th} cycles. 58

Figure 4.6 Mid column concentration profiles of bulk mobile phase protein concentration Cbi (blue), protein concentration at the outer surface of resin $Cp, R = Rp$ (orange) and the protein concentration at the center of resin particle $Cp, R = 0$ (grey) for (A) 0^{th} , (B) 50^{th} and (C) 80^{th} cycles. 59

Figure 4.7 (A) Variation of average concentration differences across pore (δp), film (δf) and overall (δ), (B) Fractional contribution of pore ($\delta p/\delta$) & film resistance ($\delta f/\delta$) towards overall driving force. 59
Figure 5.1 (A) Schematic illustration of the set up for Online-HPLC, (B) Configuration of 2 position 6 port valve for loop filling, (C) Configuration of 2 Position 6 port valve for column loading. 66
Figure 5.2 Proposed method for online analysis and pooling decision. 67
Figure 5.3 (A) Preparative CEX chromatogram showing separation of aggregates from the mAb product. The insets show online SE-HPLC chromatograms at various points of the elution profile. It is evident that aggregate species elute in the later part of the product peak, (B) Comparison of off-line and online HPLC analysis for monomer peak %. 69
Figure 5.4 (A) Interpolated graph of online data showing distribution of aggregates and monomer with elution CVs, (B) Contour plots indicating impact of elution start and end CV on % pool purity. The reddish area shows the acceptable range for pooling. 70
Figure 5.5 (A) Preparative CEX chromatogram showing separation of charge variants of the mAb product. The insets show online CE-HPLC chromatograms at various points of the elution profile, (B) Comparison of off-line and online HPLC analysis for main species peak %. 72
Figure 5.6 (A) Interpolated graph of online data showing distribution of acidic, main and basic species with elution CVs, (B) Contour plots indicating impact of elution start and end CV on % main species pool purity. The reddish area shows the acceptable range for pooling. 73
Figure 5.7 Preparative CEX chromatographic profiles for 2 column CMC operation consisting of 4 elution profiles (2 run per column). 74
Figure 5.8 (A) Plot of % purity with online fraction number from which only highlighted region was utilized for modelling, (B) Quadratic fit for % point purity model with online fraction number, (C) Quadratic fit for % pool purity with fraction purity. 76
Figure 6.1 Schematic of PLC controlled continuous integrated system coupled with online HPLC for monitoring of charge variant composition. Orange, green, grey and violet arrows indicate data streams collected from different unit operations including pumps and surge vessels and stored by PLC. According to control strategy, the new calculated set points were sent by PLC back to BioSMB for maneuvering of pumps and valves switching as shown by red arrow. The solid line (blue) shows the fluid flow path for continuous train. 84
Figure 6.2 Gantt chart indicating the scheduling of different phases of integrated unit operations of continuous mAb train. 85

Figure 6.3 Analytical CE-HPLC chromatograms for feed of different charge variant composition in the range of acidic: 13-17%, main: 18-23% and basic: 59-68%, respectively. A, M, B1 and B2 denotes acid, main, base 1 and base 2 variants respectively. The % of the various charged species (A, M, B1, B2) with their respective areas are identified on the chromatogram as mentioned throughout the manuscript. The absorbance units (UV 280) are plotted on y axis. 93

Figure 6.4 Breakthrough curve fitting for different concentrations (1.02-2.11 mg/ml) (A) Mechanistic general rate model, (B) Empirical normalized error function model, (C)-(D) Correlation of empirical model parameters (k and τ) obtained by fitting equation 6.14 with concentration.1 CV corresponds to 4 mL.....96

Figure 6.5 Comparison of the simulated elution charge variants profiles of with the experimental value. Dotted points show the experimental data. 98

Figure 6.6 Actual by predicted plots of ti (left) and tf (right) models.....99

Figure 6.7 Framework of control strategy algorithm for model-based peak cutting criteria based on concentration analytics using HPLC to achieve desired composition of charge variants in CEX eluate. 101

Figure 6.8 Time varying charge variants composition of (A) online HPLC data before and after perturbation (B-D) of SV2, CFIR, and SV3 simulated data after the perturbation was introduced before shut down. 103

Figure 6.9 (A) UV elution profiles of CEX consisting of 6 cycles and 18 runs (B) Zoomed view of an elution peak of CEX run along with pool cut points (C) Pool cut points (ti and tf) for entire continuous train. 104

List of Tables

Table 3.1 Details of PCC protein A Chromatography method for unit cycle. 26
Table 3.2 Model architecture of LSTM autoencoder with respective parameters for each layer of encoder and decoder. 36
Table 3.3 Model output in terms of R^2, RMSE, and PE (%) of experimental vs. predicted yield % of PLS, ANN and PLS-ANN model. 40
Table 4.1 Variation of Biot number with number of cycles. 55
Table 5.1 Comparison of the interpolated predicted values and the actual experimental values. 71
Table 5.2 Details of CMC CEX Chromatography Operations for Unit Cycle. 74
Table 6.1 Details of continuous Protein A and CEX chromatography operations for unit cycle under steady state conditions for protein of concentration 1 mg/mL. 3 columns were utilized for both chromatography steps where 1 CV corresponded to 4 mL. 86
Table 6.2 Optimized model parameters obtained from the inverse fit of elution profiles. 97
Table 6.3 Sorted Parameter Estimates for ti and tf models. 99
Table 6.4 Validation for 2 different feed showing predicted and actual values for ti and tf. 100

List of Abbreviations

AE	Auto-Encoder
AEX	Anion Exchange Chromatography
ANN	Artificial Neural Network
BET	Brunauer–Emmett–Teller
CADET	Chromatography Analysis and Design Toolkit
CE-HPLC	Cation Exchange High Performance Liquid Chromatography
CEX	Cation Exchange Chromatography
CFIR	Coiled Flow Inverter Reactor
CIP	Cleaning in Place
CMC	Concurrent Mode Chromatography
CP	Cleaning Procedures
CPP	Critical Process Parameters
CQA	Critical Quality Attributes
CSTR	Continuous Stirred Tank Reactor
CV	Column Volume
DBC	Dynamic Binding Capacity
DCS	Distributed Control System
DLVO	Derjaguin-Landau-Verwey-Overbeek
DoE	Design of Experiment
EKF	Extended Kalman Filter
FDA	Food and Drug Administration
FEM	Finite Element Method
FTIR	Fourier Transform Infrared
GP	Gaussian Processes
GRM	General Rate Model
HETP	Height Equivalent of a Theoretical Plate
HPLC	High Performance Liquid Chromatography
IE-HPLC	Ion Exchange High Performance Liquid Chromatography
IF	Isolation Forest
LDF	Linear Driving Force
LKM	Lumped Kinetic Model
LRM	Lumped Rate Models
LSTM	Long Short Term Memory
LV	Latent Variables
mAb	Monoclonal Antibody
MAE	Mean Block Error
MC	Mixing Cells (in Series Model)
MCD	Minimum Covariance Determinant
MCSGP	Multicolumn Counter-Current Solvent Gradient Purification
ML	Machine Learning
MSE	Mean Square Error
NIPALS	Nonlinear Iterative Partial Least Squares
NIR	Near Infrared Spectroscopy
OC	Orthogonal Collocation
OCFE	Orthogonal Collocation on Finite Elements

PAT	Process Analytical Technology
PC	Principal Component
PCA	Principal Component Analysis
PCC	Periodic Counter-Current
PCHIP	Piecewise Cubic Hermite Interpolation Polynomial
PE	Percentage Error
PFR	Plug Flow Reactor
PLC	Programmable Logic Controller
PLS	Partial Least Squares
PLS-ANN	Partial Least Squares-Artificial Neural Network
QbD	Quality by Design
RDM	Reaction Dispersive Model
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
RP-HPLC	Reverse Phase High Performance Liquid Chromatography
RTD	Residence Time Distribution
RTRT	Real Time Release Testing
SCADA	Supervisory Control and Data Acquisition
SE-HPLC	Size Exclusion High Performance Liquid Chromatography
SMA	Steric Mass Action
SMCC	Sequential Multi-Column Chromatography
SPFF	SP Sepharose Fast Flow
SV	Surge Vessel
TDM	Transport Dispersive Model
UHPLC	Ultra High Performance Liquid Chromatography
UV	Ultraviolet