TOWARDS UNCERTAINTY QUANTIFICATION OF COMPLEX CHEMICAL PROCESSES WITH APPLICATION TO POST-COMBUSTION CARBON CAPTURE

A Dissertation
Presented to
The Academic Faculty

By

Jayashree Kalyanaraman

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemical and Biomolecular Engineering

Georgia Institute of Technology

May 2017

TOWARDS UNCERTAINTY QUANTIFICATION OF COMPLEX CHEMICAL PROCESSES WITH APPLICATION TO POST-COMBUSTION CARBON CAPTURE

Approved by:

Dr. Matthew J. Realff, Advisor
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Yoshiaki Kawajiri, Co-advisor
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Christopher W. Jones
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Ryan P. Lively
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Roshan J. Vengazhiyil
Department of Industrial Systems and Engineering
Georgia Institute of Technology

Date Approved: Mar 30, 2017

And, when you want something, all the universe conspires in helping you to achieve it.

Paulo Coelho, The Alchemist

To my parents, my husband and my son

ACKNOWLEDGEMENTS

I am thankful to a lot of people, without whom I wouldn't have pursued and completed my PhD program. First of all, I would like to sincerely thank Dr. Carson Meredith, who allowed me to join in the Spring, the off-semester as the new graduate students are usually intaken only during Fall semester, when I could not register due to my visa issues. I am ever thankful to Dr. Matthew Realff for taking me into his group by converting a post-doctoral position to a PhD position. During the course of my thesis, I have had many good discussions with my fellow doctoral students, Gaurav Agarwal, and Jason Bentley and post-doctoral researchers Swernath Subramanian, and Fateme Rezeai and I am thankful to them. I am also highly thankful to Dr.Yanfang Fan and Dr.Fateme Rezeai, with whom I have had an excellent colloboration in designing experiments together and also for providing me a guided tour of their experimental facility. More imporantly, I am indebted to both my advisors, Dr. Matthew Realff and Dr. Yoshiaki Kawajiri for supporting me and allowing me to work from home, especially after my motherhood, which helped me immensely to balance my PhD and personal life. I am thankful to Dr. Christopher Jones and Dr. Ryan Lively for providing technical guidance and insight from materials perspective during the development of the carbon capture process model. I would like to thank Dr. Roshan Vengazhiyil, whose excellent teaching in his course "Advanced statistical modeling ", helped me to gain a proper statistical background, which was very useful in my research. I am very thankful for all the technical discussions that I have had with my advisors, Dr. Matthew Realff and Dr. Yoshiaki Kawajiri, which helped me tremendously both for my research as well as to improve my overall research outlook. I would like to thank all my committee members, Dr. Matthew Realff, Dr. Yoshiaki Kawajiri, Dr. Christopher Jones, Dr. Ryan Lively and Dr. Roshal Vengazhiyil for asking critical questions and providing suggestions during my review presentations. Although it is not part of my culture to thank family explicitly, I would like to acknowledge their support here so that I could remember it forever. Espe-

cially, I am thankful to my husband, Balaji Muralidharan, who have helped me throughout my PhD providing me technical suggestions and guidance at every stage, acting almost like my third advisor. I am also thankful to my parents and both my in-laws, who besides being providing encouragement, have taken the trouble in travelling away from home and staying in the United States for months to help me and my husband, in taking care of our little one, Rohan, while we try to complete our PhD. Finally, most importantly I would like to thank U.S. Department of Energy for financially supporting my research.

SUMMARY

Quantifying the extent of model uncertainty is crucial in the technical feasibility analysis of energy technologies and can provide a significant saving of cost and time. However, performing the uncertainty quantification for a complex chemical process involving coupled PDEs system is computationally prohibitive. Parallel algorithms and parallelization is utilized wherever possible in the entire framework of uncertainty quantification to handle the involved computational cost. The model complexity, on the other hand, is retained without resorting to any form of reduction or surrogate modeling. The application that is studied to perform the uncertainty analysis is the post-combustion carbon capture via Rapid Thermal Swing Adsorption using amine sorbents in a hollow fiber contactor. Thermal Swing Adsorption is a dynamic non-isothermal cyclic process with a complex interplay of mass transfer kinetics and equilibrium, and therefore the governing process model is a complex coupled system of PDEs. The process model developed is initially calibrated using conventional methods of parameter estimation and the performance is benchmarked for comparison against the results obtained incorporating uncertainties.

The computational challenge in performing Bayesian inference, is handled by employing Sequential Monte Carlo, a parallel algorithm based on particle filtering. The uncertainties involved in the process are characterized using four different approaches, viz ”Hierinf”: the data are separated into subsets and the inference is performed for the individual series, ”Varinflat-inf”: the variance of the parametric uncertainties are increased by increasing the variance of the residual errors, ”Uresvar-inf”: wherein, additional model parameters are considered as uncertain in an attempt to reduce the residual variability (errors), ”Mdiscrep-inf”: wherein, the additional uncertainty is introduced in the model structure via the model discrepancy term. The characterized uncertainties, obtained from each of the four different approaches are propagated through the process model and the uncertainties in the key prediction variables, viz: the product quality and process performance, viz:

CO_2 swing capacity are obtained. The last component of uncertainty analysis is to be able to design experiments optimally in order to reduce the prediction uncertainties. A new method is proposed wherein the prediction uncertainty is reduced through designing experiments based on the utility function formulated with the parametric distributions. The proposed method is demonstrated for a simpler system of RTSA, in which only the adsorption isotherm parameters are considered as uncertain.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION AND BACKGROUND

Mathematical modeling of physio-chemical processes is performed for various purposes; to get an insight into the process behavior at scale-up conditions, to perform optimization studies in order to maximize productivity, product quality etc, for optimal design of experiments, or to study its techno-economic feasibility. For all of the above cases, it is imperative to validate the model predictions against the experimental data, to use the model reliably and effectively. However, in validation, the model predictions will not match all the experimental data, given that there are a number of uncertainties in both the experimental data as well as in the model. In order for the model to predict the physical process robustly amidst all the uncertainties, it is necessary to incorporate the effect of all the uncertainties in the model. Thereby, the impact of the uncertainties on the model predictions can be quantified, making the model reliable. The calibration of all uncertainties and quantifying their impact on the model are together described as Uncertainty Quantification (UQ). UQ is increasingly important in the recent years across many engineering fields, especially in hydrology [1], fluid dynamics [2], systems biology [3], bio-chemistry [4] and chemical engineering [5, 6, 7].

The overarching objective of this thesis is to develop a numerical and an algorithmic framework, which can be used to perform a comprehensive UQ analysis for a complex chemical process model. The motivation and the novelty of the thesis are described in the following sections.

1.1 Motivation to perform UQ

In conventional modeling practice, the model parameters are tuned so as to match some selected experimental data [8]. The model thus obtained, based on the best fit with the experimental data, is then used for all predictive applications. However, though the model predicts well at certain experimental conditions, it may be inaccurate at other conditions due to various uncertainties, both in experimental data and in the model. Therefore the model, when used for predictive applications for different scale such as plant-scale at which the predictions are not validated with some observed data, can be inaccurate and completely unreliable. In order to overcome such shortcomings of mathematical modeling, industries typically employ multiple pilot plant demonstrations of the process in order to establish the reliability of the model and also to correct for its inaccuracy [9]. On the other hand, if the model is calibrated rigorously, incorporating all the uncertainties that are observed in the experimental data, including model inadequacy, the model predictions can be reliably used for predictive applications, perhaps avoiding or reducing a large number of such multiple scale demonstrations.

1.2 Sources of uncertainties

The uncertainties that are of concern while doing mathematical model predictions can arise from a number of sources [10] and are typical of any physio-chemical process. The uncertainties could either be due to uncorrectable errors in the observed experimental data or be present in the mathematical model itself. By error, what is meant is the difference between the true value and the measured value and when it cannot be corrected and it leads to uncertainty in the outcome. There can be various sources of errors and can be largely categorized as epistemic (systematic) or aleatory, based on the notion, if in principle (with additional measurements or knowledge) the error can be known or not.

Noise in measurement devices or random variability in experimental conditions can be

classified as aleatory. The presence of aleatory uncertainties in experimental data is largely unavoidable due to the accuracy limitations of measurement devices and the complexity of experimental system involving several manual synthesis steps, each introducing a probability of variation. For example, synthesis of hollow fiber sorbent membrane module [11] involves several manual synthesis steps. With each of the manual steps introducing some random variability, the overall properties of the hollow fiber module synthesized varies, thereby affecting the experimental outcome such as the breakthrough curve [12]. Although it is possible to reduce the aleatory uncertainty by conducting a number of repeat measurements, it may not be easy and can be very time consuming to perform repeat measurements of complex physical processes. The aleatory errors in the measurement data are carried over and incorporated in the model as uncertainties in the physical model parameters, while doing parameter estimation. In addition to the physical model parameters, there are other additional parameters called as hyper-parameters that are introduced while doing the UQ [13]. Mean and standard deviation of the errors (characterized by any probability distribution) are some examples of hyper-parameters, which are typically used to quantify or characterize the uncertainties and are estimated along with the other physical model parameters.

The epistemic uncertainties, on the other hand, arise from either incomplete measurement (lack of information) or from model inadequacy in describing the actual physics that occurs in the process. For example, various physical parameters in the model such as pore diameter, thermal conductivity, diffusivity etc. are not measured directly and are mostly estimated based on parameter estimation techniques by fitting the model with other observable experimental data. An indirect estimation, in turn, can introduce uncertainties in those parameter values, which in principle, have a precise value. Another example can be the uncertainties that arise due to randomness in hollow fiber module packing. As it is very difficult to control the placement of fibers within the module, it can be random every time the fibers are loaded in the module [12] for a new experimental run. Although it is actually

possible to determine the exact location of fibers within the module after they are placed, it is a very difficult measurement to make. Besides the incomplete measurements, the model itself may be inadequate due to the various simplifying assumptions that are made to reduce the computational complexity and thereby neglecting or approximating some physics observed in reality [10]. These epistemic uncertainty are either characterized and incorporated either as uncertainties in the physical model parameters when the model is adequate (i.e. adequately parametrized) or as a model discrepancy term if the model was found to be structurally or parametrically inadequate. By adequacy, it is meant that the model with its existing structure and the physical model parameters is able to explain the observed experimental data within the assumed error range. Figure 1.1 illustrates the various sources of uncertainties that are described and how they can affect the model predictions.

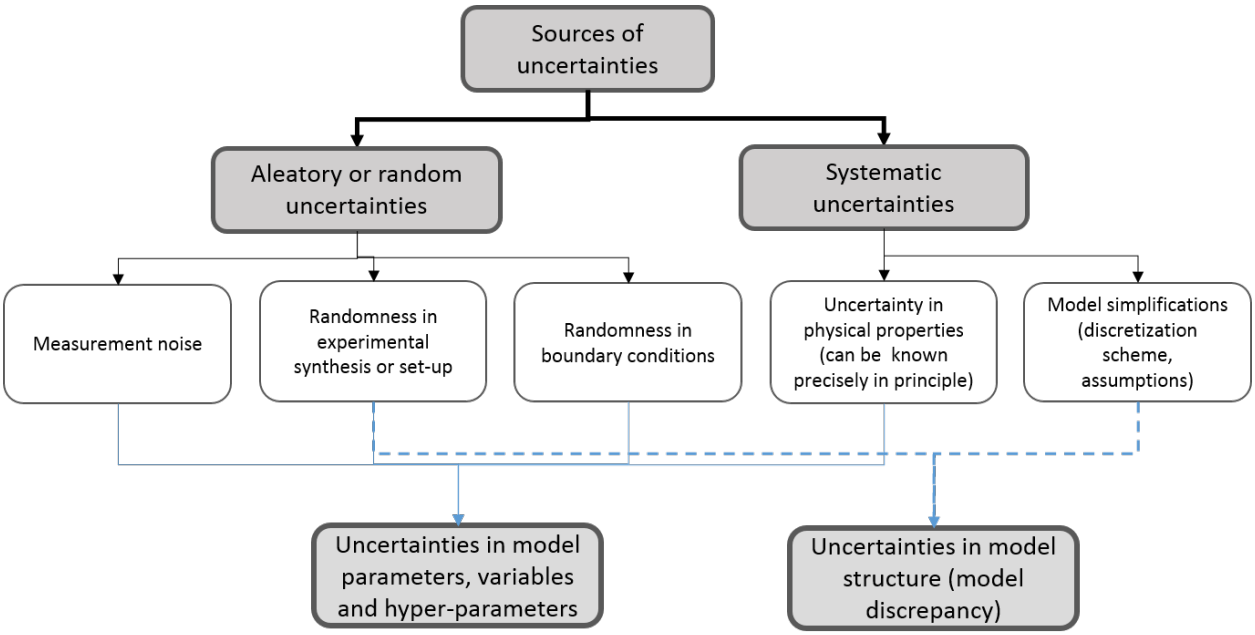


Figure 1.1: Various sources of uncertainties in the mathematical model

1.3 Bayesian analysis of UQ

In a probabilistic framework, Bayesian analysis based UQ is the most commonly used technique to characterize and quantify the uncertainties on model predictions [14]. Bayesian inference offers several advantages over conventional (frequentist) methods such as maximum likelihood estimation [8], boot strapping etc [15]. Primarily, it does not require modification to the model (such as linearization) and provides a comprehensive treatment of parametric and model uncertainties. In contrast to the maximum likelihood estimation, which assumes that the model mismatch errors are normally distributed, Bayesian method determines the exact distribution of the uncertain parameters commensurate with the evidence of experimental data. Besides, Bayesian method also has the provision of utilizing the prior knowledge of the parameters while doing the inference, which can be very useful in cases where experimental data are sparsely available and have huge uncertainties. It also facilitates sequential addition of experimental data, while determining the probability distribution of parameters (called the posterior parametric distribution). The Sequential Bayesian Update (SBU) is of tremendous significance [13], in cases where collecting the experimental data is time consuming and also when the model evaluation is computationally intensive as is the case of complex chemical processes. By sequentially updating the parametric probability distribution with the experimental data one at a time, computational cost of inference can be reduced, as will be described in Chapter 4. Also, the sequential Bayesian update is very impactful in the design of experiments, wherein the latest parametric distribution needs to be updated with only the newly gathered experimental data instead of determining the parametric distribution using all the experimental data. A brief description of each of the components of UQ analysis under Bayesian framework, viz. characterization of uncertainties, propagation of uncertainties through the model and design of experiments to reduce uncertainties, is described in the following subsections.

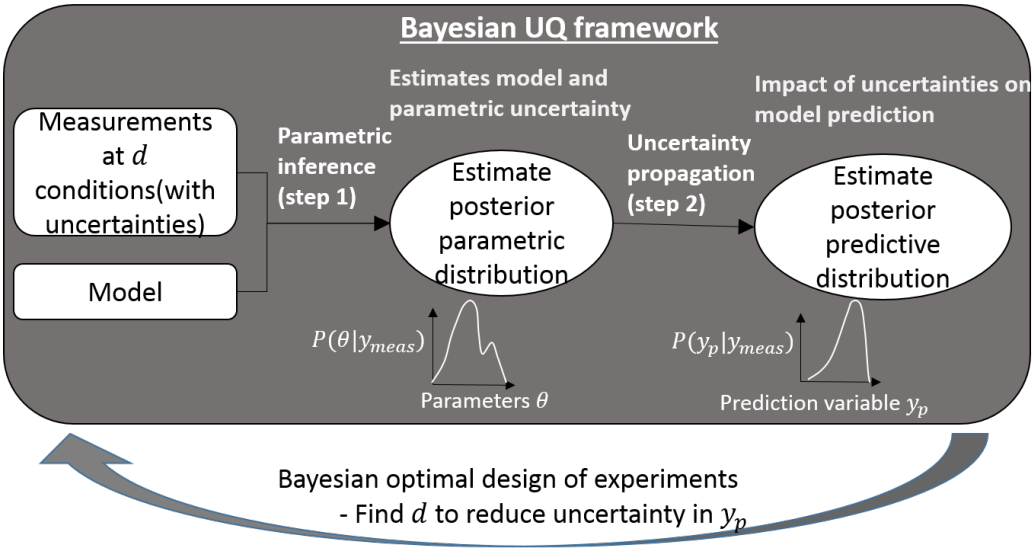


Figure 1.2: UQ framework under Bayesian analysis

1.3.1 Characterization and propagation of uncertainties

As the first step of UQ, calibration or characterization of uncertainties is performed to yield a joint probability distribution of all the parameters (both model parameters and hyper-parameters characterizing the distribution) based on the observed experimental data [14]. The parametric probability distribution, thereby obtained, is commonly known as posterior parametric distribution, as it is obtained posterior to observing the experimental data. The next step of the UQ analysis is to propagate the obtained parametric probability distribution through the process model to determine the probability distribution of the observed variables that are of key interest such as product quality (purity and recovery), product throughput, total operating or capital cost of the process. Figure 1.2 illustrates the general concept of UQ under Bayesian framework. In contrast to the conventional modeling approach, which relies on the model prediction that are based on the ideal conditions (i.e. without any influence of uncertainties), the UQ analysis provides a complete distribution of the model predictions incorporating the effect of all the uncertainties. Thereby, it enables definition of the worst and best-case scenarios, in terms of either product performance or

process economics metrics that can be encountered in reality within certain probability bounds. The worst case scenario, here is defined as the worst outcome that can occur with a certain probability, which is considered as 99% in this work. Similarly, the best case scenario is defined as the best outcome that can occur with the same probability.

1.3.2 Design of experiments to reduce uncertainties

Apart from characterizing and quantifying the uncertainties and their impact on model predictions, Bayesian based UQ analysis can also be used to determine the optimal design of experiments to gather data and thereby reduce the uncertainties to the desired level [16]. In most chemical engineering processes, obtaining experimental data at a large number of operating process conditions may be very expensive and time consuming. Therefore, under such scenarios, it is advantageous to design the experiments optimally so as to gain maximum information with the least amount of experimental data, which is termed as optimal experimental design (OED) [17].

In the frequentist approach, OED is determined by applying the well known *alphabetical optimality criteria* ([18] with the Fisher Information Matrix (FIM) and has been successfully applied for models with linear parametric dependence. In case of models with nonlinear parametric dependence, a number of simplifying assumptions are made to employ FIM based optimal experimental design, including linearization of the model response and Gaussian approximation of the parametric distributions [19]. Bayesian experimental design, on the other hand, does not require, in general, [16] any simplifying assumptions on the parametric distribution or model linearization. Measure of information gain using Bayesian analysis, is represented by an objective function called the *utility function*, which includes an optimality criteria and is maximised to determine the optimal design condition. A detailed review on the Bayesian experimental design methods and the variations of the utility functions can be found in Chaloner and Verdinelli (2009) [16].

The existing literature on applying UQ for chemical processes have largely been ei-

ther for batch process models such as reaction kinetics models [20, 21, 5], molecular dynamic models [22], adsorption isotherm models [23, 24] involving several ODES and only very few studies have been performed for large models [packed bed adsorber models [25], crystallization [6], fluidization bed [26]] involving PDEs. Similarly, the studies on applying Bayesian inference to optimal design have been limited to simple linear models. The bottleneck in applying Bayesian UQ analysis for complex large scale models is the computational cost involved, which can get prohibitive and will be detailed in the following sub-section.

1.4 Computational challenge with UQ analysis

Though there are several advantages in performing UQ analysis of mathematical model predictions, it also has the downside of presenting a huge computational challenge when it comes to employing it for actual chemical process models. The main computational expense in doing the UQ is in the inference step of UQ, which requires numerous evaluations of the model. The inference problem, which determines the posterior probability distribution of parameters, typically requires about 10^4 - 10^5 simulations when done using conventional simulation methods such as Markov Chain Monte Carlo (MCMC). Hence, it becomes prohibitive to deploy the MCMC with the complete process model.

Most of the studies on applying Bayesian inference for large chemical process models (involving coupled PDEs) and large number of parameters have employed different methods of surrogate or reduced order modeling to reduce the associated computational cost. A number of reduced order modeling techniques are being studied widely to perform UQ of complex large scale models. Among them, spectral methods such as stochastic Galerkin [27] and stochastic collocation [28] are widely applied, mostly for the uncertainty propagation step of UQ. While the former is an intrusive method which involves transformation of the complete process model by Galerkin projection with polynomial chaos expansion, the latter is a non-intrusive method and does not involve any model transformation. Besides,

Proper Orthogonal Decomposition (POD) [29] and Kriging models ([30]) are the other commonly used methods to obtain surrogate models [27] and Kriging approach. Although surrogate models, once determined, can reduce the computational cost of the inference problem significantly, determining a sufficiently accurate reduced order model is a significant challenge and can be very cumbersome and computationally quite intensive.

In case of Bayesian experimental design, the computational expense of evaluating the utility function (the objective function of optimization) has been a major hurdle in deploying the Bayesian design to determine the optimal experiment as most of the real world models are complex and cannot be analytically evaluated [31], [32]. In effect, most of the reported work on Bayesian experimental design have either been using linear models and in the few studies having non-linear models, approximations of the utility function or Gaussian approximations of the posterior distributions are used [33]. In that context, Muller and Parmigiani [34] suggested using a Monte Carlo estimator and simulation based optimal design by fitting the Monte Carlo samples of the utility surface. However, they concluded that the evaluation becomes computationally prohibitive for large dimensions of design parameters. In recent work, Solonen *et al.* [35] applied the simulation based optimal design for a CSTR model using variance of predictions as the utility function. They sidestepped the computational complexity of posterior distribution evaluation after every added experiment, by weighing the parameters with the corresponding likelihood of the new measurements. Even with such methodology, the likelihood evaluation could turn out be computationally expensive for complex models. More recently, Huan *et al.* [36] applied Bayesian D-optimality for experimental design involving a combustion kinetic model. They reduced the computational expense of likelihood evaluation by using a polynomial chaos surrogate.

Of the two routes to tackle the computational complexity of the inference problem, the use of surrogate models reduces the overall cost by reducing the cost of a single simulation and thereby providing the overall speed-up. The alternative route, however, is to use full

models without reduction and rather use some algorithmic techniques to gain speed-ups of the overall cost. For example, the use of transport maps from normal distribution to the actual posterior distribution to accelerate MCMC, or the use of parallelism in the computation with the help of parallel algorithms such as Sequential Monte Carlo (SMC) [7], also known as Transitional MCMC [37] (T-MCMC). Although, there exists many other efficient sampling techniques such as Latin Hypercube Sampling (LHS) [38], Quasi Monte Carlo (QMC) [39], etc. besides MCMC and SMC, they can be applied only when the parametric distribution (the posterior probability distribution) is known as in the case of uncertainty propagation step, in which samples from the parametric distribution are drawn and propagated through the model. The inference problem, to the contrary, is an inverse problem where the distribution itself is unknown and needs to be determined given the experimental data.

In this thesis, the involved computational complexity in applying the Bayesian UQ analysis for a complex chemical process model is tackled using parallel computation, through the development and implementation of parallelizable algorithms for all the steps of UQ analysis.

1.5 Application: Post-combustion carbon capture via adsorption in hollow fiber sorbents

The complex chemical process that is of interest in this thesis is post-combustion carbon capture using hollow fiber based solid sorbents employing rapid thermal swing adsorption process.

Post-combustion CO_2 capture is the route typically followed for traditional coal fired power plants as it allows an easy retrofit into the existing infrastructure [40]. Among the different methods like chemical looping combustion, liquid absorption and membrane processes that can be used for post-combustion CO_2 capture, adsorption on microporous materials, utilizing a temperature or pressure swing regeneration, has been found to be a poten-

tially energy efficient technology compared to liquid absorption or membrane processes. CO_2 removal by adsorption process is promising in terms of relatively high energy efficiency and is being investigated extensively in the community with several variations in the operating configurations such as vacuum, pressure or temperature swing, and also with varied sorbents [41]. However, conventional cyclic swing adsorption process in packed bed design suffers from the high economic cost of pressuring the flue gas in large scale for pressure swing and longer regeneration time for temperature swing processes.

In recent work, Lively et al [42] demonstrated a novel hollow fiber based solid sorbent system for CO_2 capture which overcomes the deficiencies of conventional cyclic swing adsorption processes. The technology uses a polymeric hollow fiber, shown in Fig.2.3, loaded with sorbent to be used in a Rapid Thermal Swing Adsorption (RTSA) process. The bore side of fiber is coated with a dense impermeable lumen layer. Such a fiber morphology allows the flow of cooling water or steam through the bore side without permeating through the fiber walls, thus making the hollow fiber act as a micro heat exchanger. As the actual feed flue gas stream from the power plant to CO_2 capture unit would be saturated with water, the microporous hollow fibers are loaded with supported amine sorbents which are highly efficient CO_2 adsorbents even under humid conditions.

CO_2 capture using RTSA is a fairly complex chemical process and the model that describes the dynamic non-isothermal adsorption/desorption phenomenon in the hollow fiber bed involves a set of highly nonlinear coupled PDEs. Each simulation of a single step (for eg.adsorption) requires upto 5 minutes of CPU time for integration with Python using an AMD 2.4GHz processor. Hence, it is impossible to perform Bayesian inference with traditional methods such as MCMC without use of some techniques to tackle the computational challenge. This therefore leads to the objectives of this thesis, which are listed in the next subsection.

1.6 Research objectives and thesis organization

The overarching objective of the thesis is to develop a numerical and an algorithmic framework which can be used to perform a comprehensive UQ analysis of a complex chemical process model. The term UQ analysis, refers to both characterization of uncertainties using Bayesian inference as well as propagation of the characterized uncertainties through the chemical process model to determine their impact on the observed variables of interest at the desired process operating conditions. Additionally, this thesis also aims to provide a methodology to reduce the uncertainties in the model predictions via optimal design of experiments. The model complexity is retained without resorting to any form of model order reduction, however, parallel algorithmic techniques are used in all the steps of UQ analysis to reduce the computational overhead. The specific application for which the developed framework is demonstrated and applied is the post-combustion carbon capture using a rapid thermal swing adsorption process carried out with hollow fiber sorbent modules. The overall objective is accordingly sub-divided as the following five components, each of which is described in detail as a separate chapter.

1.6.1 Development of the process model

As part of this sub-task, rigorous mathematical models are developed to describe the cyclic non-isothermal process of rapid thermal swing adsorption with hollow fiber adsorbents. Detailed models describing the complex coupled mass and heat transfer kinetics are developed based on first-principles [43]. Cyclic simulation is performed and tested using some arbitrary values for physical model parameters (from literature) and the key process design variables of interest are determined [44].

1.6.2 Conventional parameter estimation and its limitations

The unknown key model parameters are determined by fitting the model with the experimental data using Maximum Likelihood Estimation (MLE). The model thus established under the conventional setting of parameter estimation, is then tested for its prediction robustness against varied experimental conditions, such that the data encompasses all the aleatory and systemic errors. Thereby, the motivation to perform UQ analysis for this mathematical mode is presented [43].

1.6.3 Characterization of uncertainties: algorithm and computational framework development

The uncertainty characterization is performed using Bayesian inference and the full model without reducing its complexity. Therefore, the computational cost involved in performing Bayesian inference for the complex process model is is handled by employing Sequential Monte Carlo (SMC), a completely parallel particle filter based algorithm. A complete scalable framework to perform Bayesian inference in parallel with adaptive SMC is developed in Python. The developed numerical code is completely modular and can be applicable for any chemical process model by replacing the carbon capture process model with an appropriate model. [7]

1.6.4 Uncertainty propagation: algorithm and computational framework development

The underlying idea of this thesis is to exploit parallel computation in UQ analysis, wherever possible to reduce the computational overhead. In the same view, the uncertainty propagation is also performed in parallel using Monte Carlo sampling of the obtained posterior parametric distribution. The numerical framework to perform propagation is developed in Python. The process design variables for which the uncertainty in predictions are of interest are the final product purity and recovery from cyclic simulations of the process. Therefore,

the propagation step is also carried out through the cyclic process model and the uncertainty in product purity and recovery are recorded at cyclic steady state conditions that are attained after the process is run for several cycles.

1.6.5 Optimal design of experiments to reduce uncertainties: algorithm and computational framework development

Bayesian design of experiments is performed to identify experimental conditions, which, when observed, can provide maximum information in reducing the prediction uncertainty of adsorption breakthrough capacity. Parallel computation of the objective function, i.e the utility function, based on simulation based experimental design, is employed to identify the optimal experimental condition. Again, the numerical framework to perform experimental design are implemented in Python [45].

1.7 Thesis contributions

This thesis advances the state of the art in the uncertainty quantification of complex chemical processes. A detailed first principles based model, which adequately describes the complex interplay of mass transfer kinetics and equilibrium effects that is observed in the physical phenomenon of adsorption on amine sorbents, is developed. The model is rigorously validated under various process and operating conditions and is found to predict the experimental behavior well. As described in the literature review, there has been very limited work in the literature on performing characterization of uncertainties for complex physical process without model reduction and there has been none on applying SMC for a complex chemical process model such as that of rapid thermal swing adsorption, the major contribution of this thesis. The entire framework of SMC algorithm implementation is done from the scratch using non-proprietary codes and is built as a completely scalable generic framework, which can be applied for any chemical process model and not specific to the application discussed in the thesis. The process model is initially developed in

gPROMS, a commercial modeling software. However, as parallel computation of process simulations are required for both characterization and propagation steps in UQ analysis and since gPROMS models cannot be easily parallelized, the entire model is re-implemented in Python with explicit discretization schemes of partial differential equations. A new formulation in terms of utility function evaluation, that can be used to reduce uncertainties in prediction based on parametric distribution evaluation alone, is proposed and demonstrated. Again, the numerical framework is implemented in Python for utility function evaluation and is easily scalable to other process models.

CHAPTER 2

DEVELOPMENT OF THE PROCESS MODEL

The focus of this thesis is post-combustion carbon capture via Rapid Thermal Swing Adsorption (RTSA) using a hollow fiber contactor. Carbon capture using RTSA with hollow fiber contactors has been proposed recently and has been studied extensively using both experimental as well as modeling, optimization and techno-economic analyses. A sorbent loaded hollow fiber, shown in Figure 2.1, is the novel component of this process. The fiber serves as an ideal platform for the efficient transfer of heat to/from the adsorption/desorption process through the flow of cooling/hot water in the fiber bore/hollow side. The porous polymer is a highly porous substrate in which an adsorbent may be chemically grafted or mechanically entrapped and provides minimal resistances to mass transfer. A number of such hollow fibers are bundled together to form a hollow fiber module, as shown in Figure 2.2. Each module resembles a micro heat exchanger with the flow of flue gas in the shell side and cooling water in the tube or the bore side of fibers.

As described in the works of Labreche et al [11] and Rezeai et al [47], there are two major classes of supported amine materials that might be used for the hollow fiber configuration. Bollini *et al* [48] and Rezeai *et al* [47] have studied the CO_2 adsorption on hollow fiber supported amine sorbents known as *Class2* fibers, in which the hollow fiber is grafted with amine functionalized silica, specifically 3-aminopropyltrimethoxysilane (APS) in the reported work. The other class of supported amine sorbents, namely *Class1*, are synthesized by physically impregnating the amine, such as poly(ethyleneimine) (PEI) on to the silica support. In recent work, Fan *et al* [12], reported the feasibility of CO_2 adsorption

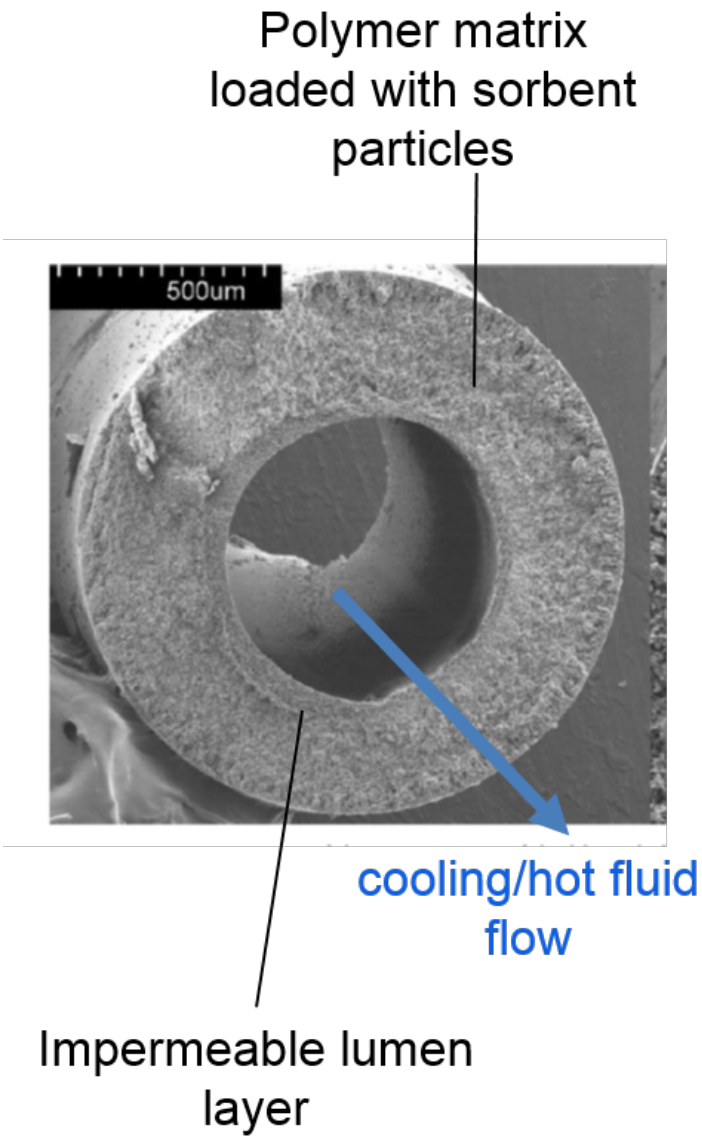


Figure 2.1: SEM of a single hollow fiber [46]

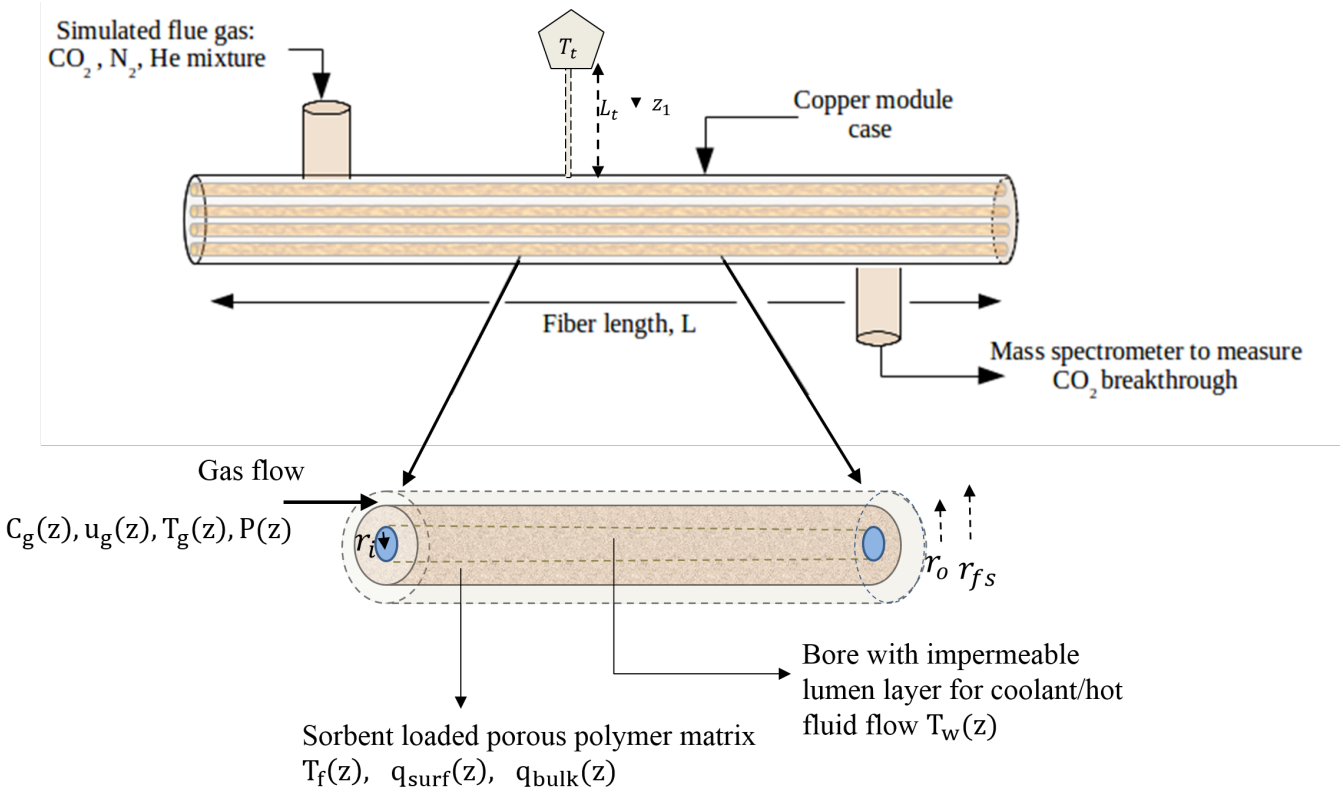


Figure 2.2: Schematic of a hollow fiber module experimental set-up

on hollow fibers post infused with *Class1* amine sorbents. In this thesis, hollow fiber sorbents impregnated with *Class1* amine sorbents is considered for modeling, data validation and UQ characterization and the data from *Class2* sorbents are used in Bayesian design of experiments. With regard to applying the model and the UQ framework for hollow fibers with sorbents other than amines, the mass transfer resistance model needs to be modified accordingly as the developed mass transfer resistance model. In particular, the micropore diffusion model may be specific to amine sorbents and not directly applicable to other sorbent materials.

Fig.2.3 shows a schematic of the RTSA cycle operation. Each fiber module shown in the figure is composed of a number of fibers bundled together. Four different fiber modules are considered to be operating in cyclic mode with each module undergoing the different steps of RTSA process in phase with others as described by Lively et al [42], [44]. A cycle starts with an adsorption step in which CO_2 containing flue gas flows in the shell side of fibers is adsorbed while cooling water flows through the bore side of the fibers, removing the released heat of adsorption. At the end of adsorption step, the module is closed at the feed end of the module and heated with hot water pumped through the bore side. This releases CO_2 creates a high pressure zone forcing CO_2 downstream driven by the developed pressure gradient and is therefore known as the self-sweeping step. At the end of the self-sweeping step, the remaining CO_2 in the shell side is swept downstream by flowing N_2 in the shell side during the N_2 purge step. Finally, during the cooling step, the temperature of fibers is returned to adsorption temperature by cooling the bed back to its initial temperature by flowing the cooling water. The RTSA of CO_2 from flue gas is a complex chemical process which is dynamic and exhibits a steady state operation only across cycles. In addition, the process operates under non isothermal conditions and has sharp adsorption and desorption fronts.

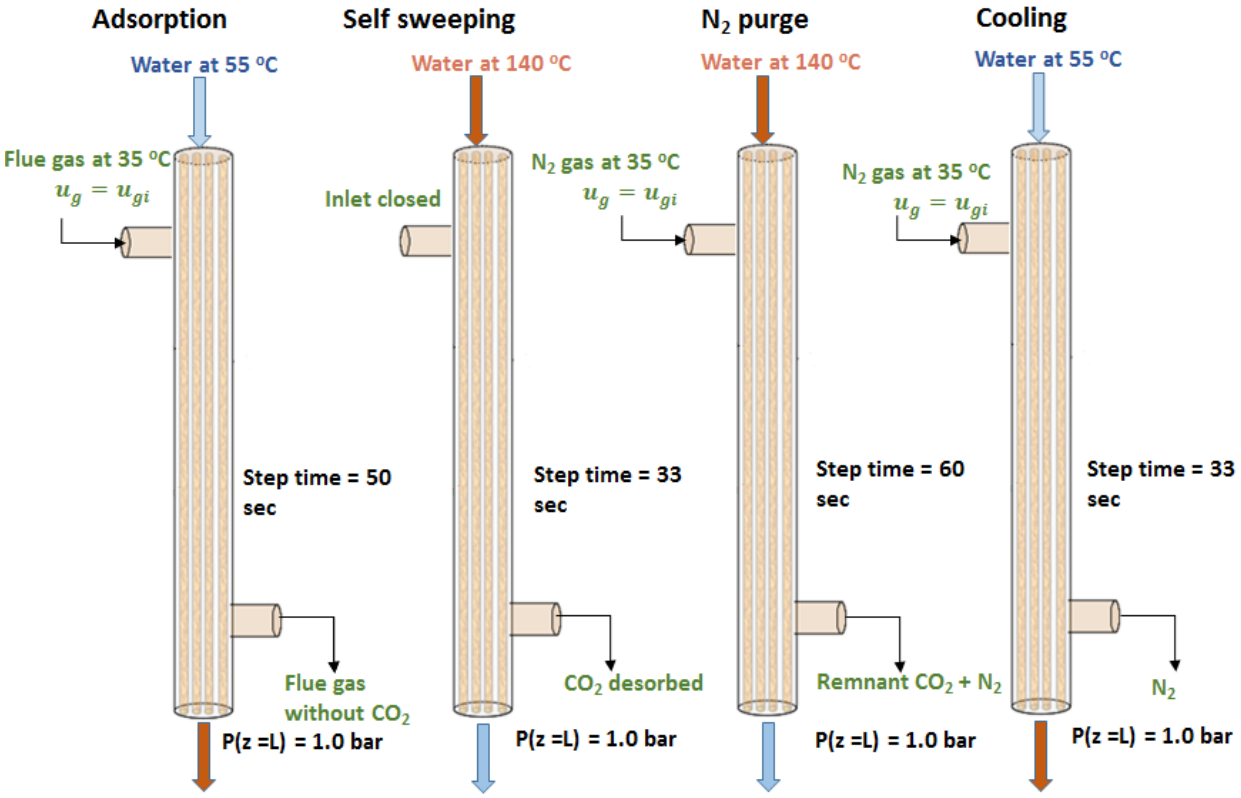


Figure 2.3: Schematic of the Rapid Thermal Swing Adsorption cycle with process stream conditions and process step times that are used throughout the work.

2.1 Experimental set-up

Experimental data that have been used in this dissertation have been collected at Georgia Tech, by collaborators of Prof.William Koros's group and Prof. Christopher Jones's group, who also worked on the project of post-combustion CO_2 capture. The details of hollow fiber sorbents spinning and amine functionalization method can be found in a previous work [11]. The fiber sorbent is a hybrid matrix composed of cellulose acetate (CA) (Sigma-Aldrich) and commercial silica particles (C803, W.R.Grace or ES757, PQ corporation) with silica loading of 57 wt$\%$. The experimental setup of the RTSA test station is shown in Fig.2.2. Unless otherwise mentioned, the ambient temperature of the test setup is maintained at 35^o C and 1 atm. Inlet flue gas entering the setup is of simulated dry gas composition containing 14$\%$ He, 14 $\%$ CO_2 and the rest N_2. All the CO_2 experiments mentioned in this thesis are conducted with no flow of cooling water through bore of the fiber [49]. The hollow fiber radial dimensions for all the experimental runs are kept fixed at $r_i = 275\ \mu m$ and $r_o = 652.5\ \mu m$ and the hollow fiber module is of internal diameter 0.15 in.

Prior to the start of each CO_2 breakthrough experiment, the module is completely purged with N_2 flow at 80 mL/min for 30 minutes at 90^o C to ensure removal of remnant CO_2 or H_2O in the system. CO_2 breakthrough data were collected at varying experimental conditions of temperature, T_{mod}, fiber length, L, and flue gas flow rates, $Q_{fluegas}$. The dead volume time in the CO_2 breakthrough curve are accounted for by subtracting the breakthrough profile of He, which is used as a tracer. The breakthrough capacity of CO_2 is calculated as the area between the He and CO_2 curve till breakthrough time t_b. Here breakthrough time t_b is defined as the time instant at which $C_{g,co_2}|_{z=L} = 0.05C_0$.

In addition to the measurements of CO_2 breakthrough data, which determines the kinetics of the process, equilibrium measurements of adsorption capacities were also obtained and used for the model validation and UQ analysis. Thermogravimetric Analysis (TGA)

experiments were conducted, to determine the equilibrium adsorption capacities of amine loaded fibers. This data is used in a separate inference/parameter estimation problem to infer only the adsorption isotherm parameters, which are later used as prior distributions while inferring the entire set of parameters from the CO_2 breakthrough data.

It is to be noted that though the experiments are not conducted with the water flow through the bores, the modeling of the cyclic process is performed including the water flow through the fiber bores.

2.2 Model Formulation

As shown in Fig.2.2, a hollow fiber sorbent module is comprised of a number of identical fibers housed within the module resembling a shell and tube heat exchanger. Adopting the approach of Happel [50] in his treatment of flow parallel to an array of cylinders, the hollow fiber sorbent module is modeled by assuming that each fiber is surrounded by a gas shell with the hypothetical boundary defined as Happel's free surface radius, d_{fs} as shown in Fig.2.2. Accordingly, the heat transfer to the fibers across the module wall is assumed to be across the hypothetical boundary for each fiber. Thus, modeling a single fiber can be used to understand the module behavior with the underlying assumption that the fibers are identically aligned against one another as rigid cylinders in a module [44]. The other basic assumptions underlying the model formulation are listed as follows.

- Ideal gas law is assumed for the gas phase.
- Radial gradients of both temperature and concentration in the fiber phase and gas phase are neglected assuming relatively negligible mass and heat transfer resistance along the radial dimension.
- N_2 and He are the non adsorbing components in the gas.

2.2.1 Mass balance equations

Mass balance equations for each of the components, viz. CO_2, He and N_2 flowing in the shell side of the fiber defined by the boundary $r\epsilon(r_o, r_{fs})$ is given by:

$$\frac{\partial C_{g,i}}{\partial t} + \frac{\partial (u_g C_{g,i})}{\partial z} - D_{ax}\frac{\partial^2 C_{g,i}}{\partial z^2} + S_i = 0 \tag{2.1}$$

$$\tag{2.2}$$

in which the indices i refer to the components CO_2, He, and N_2 respectively. In the above equation, D_{ax} is the axial dispersion coefficient which is included to account for the effect of shear flow dynamics on concentration diffusion along the axial direction and S_i is the source term denoting the rate of addition or removal of i^{th} component from the bulk gas phase. mole fraction of the inlet flue gas, P is the pressure and T_g is the temperature of the bulk gas phase. The source terms for each of the components are modeled as follows.

$$S_{CO_2} = \frac{\partial q}{\partial t}\rho_f(1-\epsilon_f)\frac{r_o^2 - r_i^2}{r_{fs}^2 - r_o^2} \tag{2.3}$$

$$S_{N_2} = 0 \tag{2.4}$$

$$S_{He} = 0 \tag{2.5}$$

The mass balance of the hollow fiber sorbent phase and the mass transfer resistance model are described in detail in the following subsection.

2.2.2 Energy balance equations

In the lab scale experimental set-up, the hollow fiber is of smaller size of length 10 to 17 inch and the module contains only few fibers. The total heat of adsorption involved, even on complete sorbent saturation, is relatively small (approximately $65J$). In effect, the adsorption heat gets rapidly lost to the module heat capacity and possibly to the thermocouple heat capacity, which are assumed to the major heat sinks in the system. These heat losses

are modeled in detail, in order to be able to match the thermal profile together with the concentration profile, while estimating the mass transfer parameters. The models for heat losses, are however, not relevant for the simulations of plant-scale fiber.

Adsorption heat loss- Experimental scale fibers

The mode of heat loss to the module is via convection from the fiber to gas and then from gas to the module. The other heat loss component which is from fiber to thermocouple is via conduction between the module and the thermocouple. This heat loss source is modeled by using an overall heat transfer coefficient for the interface between thermocouple and module, h_t. There could be perhaps some other modes of heat loss, such as direct conduction between the fibers wherever they are in contact with one another, or by conduction to the ends of the module. These, however, are not included in the heat transfer model, to avoid unnecessary complexity in the model and due to the fact that the current formulation adequately describes the experimental observations.

Components of system energy balance

Energy balance of the fiber is given in Eqn.(2.6),

$$\rho_f C_{p,f}(1-\epsilon_f)\frac{\partial T_f}{\partial t} - \lambda_f \frac{\partial^2 T_f}{\partial z^2} + \frac{\partial q}{\partial t}(-\Delta H_{ads})(1-\epsilon_f)\rho_f - 2h_g \frac{r_o}{r_o^2 - r_i^2}(T_f - T_g) + h_t(T_f|_{z=z_t} - T_t|_{z_1=0})\frac{A_t}{(r_o^2 - r_i^2)L} = 0 \tag{2.6}$$

where ΔH_{ads} is the average isosteric heat of adsorption. The axial boundary conditions of the fiber temperature, T_f are given by $\partial T_f/\partial z|_{z=0} = \partial T_f/\partial z|_{z=L} = 0$.

Heat transfer within the thermocouple needle is modeled using a one dimensional heat

diffusion equation along the dimension z_1 as given by the following equation:

$$\rho_t C_{p,t} \frac{\partial T_t}{\partial t} - \lambda_t \frac{\partial^2 T_t}{\partial z_1^2} = 0 \tag{2.7}$$

In this model, the heat from the fiber is conducted via the tip of the thermocouple needle where it touches the fiber. Fig.2.1 shows a schematic of the representation of the thermocouple along with the other relevant thermocouple parameters. The boundary conditions of the thermocouple temperature are as follows.

$$-\lambda_t \frac{\partial T_t}{\partial z_1} = h_t (T_f|_{z=z_t} - T_t|_{z_1=0}) \tag{2.8}$$

$$T_t|_{z_1=L_t} = T_{amb} \tag{2.9}$$

Here, h_t is the overall heat transfer coefficient of the thermocouple and the module and A_t is the cross section area of the thermocouple tip (Fig.2.1). The end of the thermocouple is exposed to the ambient conditions of the experimental test station which is referred to as T_{amb}.

The energy balance of the gas phase is given as follows.

$$\rho_g C_{p,g} \frac{\partial T_g}{\partial t} + \frac{\partial (T_g u_g)}{\partial z} - \lambda_g \frac{\partial^2 T_g}{\partial z^2} + h_g \frac{2 r_o}{r_{fs}^2 - r_o^2} (T_g - T_f) \tag{2.10}$$

$$+ U \frac{2 r_{fs}}{r_{fs}^2 - r_o^2} (T_g - T_{amb}) = 0 \tag{2.11}$$

where h_g is the gas convective heat transfer coefficient calculated using the Chilton Colburn analogy [51].

Water flows through the bore of the fiber to remove the heat during adsorption and to supply heat during desorption. The experiments are conducted without the flow of cooling water, and therefore the energy balance of water is used in the simulations of plant-scale

model alone. The energy balance of water flowing in the fiber bore is given as follows.

$$\rho_g C_{p,w} \frac{\partial T_w}{\partial t} + u_w \frac{\partial T_w}{\partial z} + h_w \frac{2}{r_i}(T_w - T_f) \quad = \quad 0 \tag{2.12}$$

where h_w is the gas convective heat transfer coefficient.

2.2.3 Momentum balance equations

The momentum balance of the bulk gas phase is modeled assuming that the momentum of the gas adsorbed is lost along with loss of its mass from the bulk phase. The gas phase momentum balance as given below, is same as Happel's model [50] which considers an incompressible flow through the shell side of a set of cylinders.

$$u_g(z) = -\frac{dP}{dz}\frac{1}{\mu_g}\Big(\frac{(r_o^2 - 3r_{fs}^2)}{8} + \frac{r_{fs}^4}{2(r_{fs}^2 - r_o^2)} ln\big(\frac{r_{fs}}{r_o}\big)\Big) \tag{2.13}$$

In this study, however, the velocity is not constant along the axial dimension, as the pressure drop varies with time during the time frame of adsorption. The pressure gradient is calculated using ideal gas law as follows.

$$\frac{dP}{dz} \quad = \quad R_g \sum_{i=CO_2,He,N_2} \Big(T_g \frac{\partial C_{g,i}}{\partial z} + C_{g,i} \frac{\partial T_g}{\partial z}\Big) \tag{2.14}$$

2.2.4 Mass transfer resistance model

Mass transfer kinetics have been shown to have a significant impact on the adsorption capacity of CO_2 on solid adsorbents[12] [52]. Without the use of a detailed mass transfer resistance model, it is difficult to quantify and understand the different components of diffusion resistances involved in a CO_2 adsorption process. The model developed should be such that it predicts the experimental behavior under different operating conditions and process design parameters in order to apply it for process scalability and technology fea-

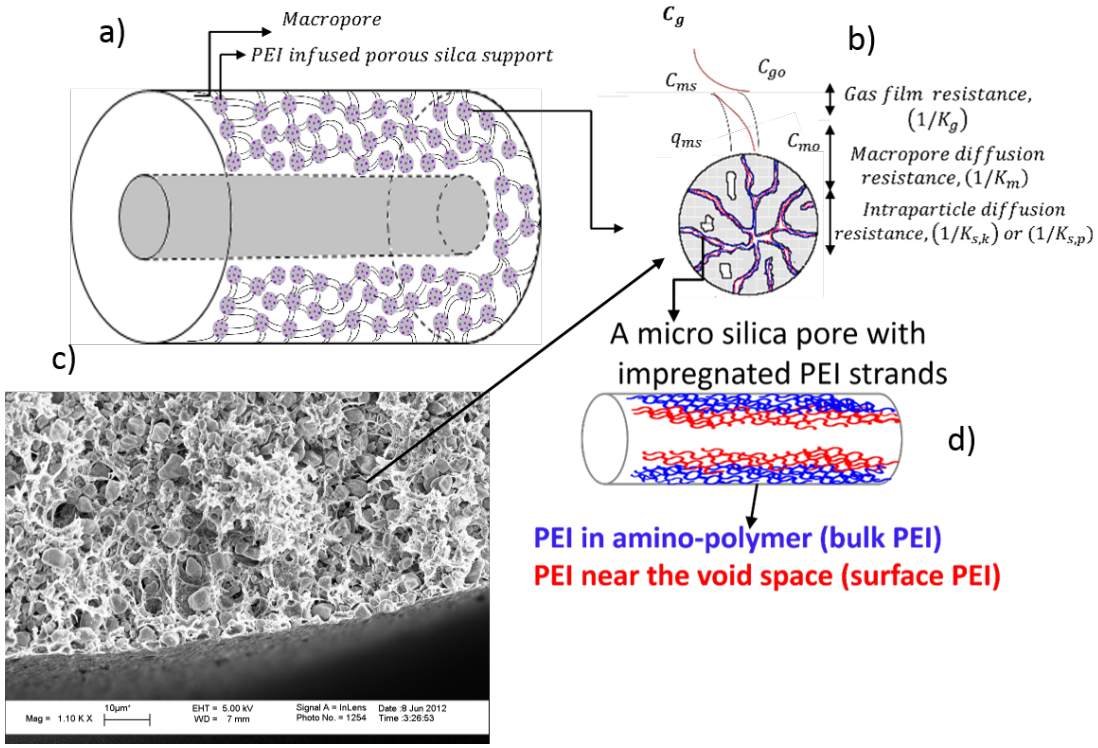


Figure 2.4: Schematic of the mass transfer resistance model a) hollow fiber with layers of mass transfer resistance, b) sorbent with micropore diffusion c) hollow fiber SEM image showing sorbents within the macropores d) a single pore schematic illustrating two regions of amine adsorption sites

sibility studies. CO_2 diffusion rather than reaction is assumed as the rate limiting step in the chemisorption of CO_2 on poly(ethyleneimine) sorbent. Conventional mass transfer resistance models used for packed beds have been unable to describe the CO_2 breakthrough from amine sorbents, which has an asymmetric shape with a sharp breakthrough and a long tail [53],[52], [54], [44]. A modification to the conventional mass transfer resistance model is thus required. It is assumed that the adsorption sites are available in two regions within the sorbent, including the ones which are easily accessible being on the surface of the sorbent pores (surface PEI) and the other which are more confined within the aminopolymer phase in the silica particle (bulk PEI). This assumption is similar to the hypothesis in Wang *et al*[55] and Bollini *et al* [52]. Accordingly, the mode of diffusion is assumed to be Knudsen diffusion (as mean free path, λ of $CO_2 = 100nm > d_{s,pore} = 20nm$) to the surface amine sites and hindered polymer diffusion to the sites in amino-polymer phase. Fig.2.4 schematically depicts the two different amine sites, as described which could be occurring in the PEI impregnated mesoporous silica.

The mass balance of the sorbent phase is modeled using Linear Driving Force model (LDF) as shown in the following equations,

$$\frac{\partial q_{surf}}{\partial t} = K_{ov,surf}(q_{eq}\psi - q_{surf}) \tag{2.15}$$

$$\frac{\partial q_{bulk}}{\partial t} = K_{ov,bulk}(q_{eq}(1-\psi) - q_{bulk}) \tag{2.16}$$

$$q = q_{surf} + q_{bulk} \tag{2.17}$$

where q_{surf} and q_{bulk} refer to the CO_2 concentration in the surface sites and in the amino-polymer sites respectively. Here, ψ is the ratio of the adsorption sites available on the surface PEI to the total PEI and is defined as follows.

$$\psi = \frac{N_s}{N_t} \tag{2.18}$$

where N_s is the number of amine sites available on the surface of the sorbent pores per unit weight of sorbent and N_t refers to the total number of amine sites available per unit weight of sorbent.

Since adsorption is assumed to occur in two different sites each with different modes of diffusion, separate mass balance equations are written for each of the sites with the respective overall mass transfer coefficient and the corresponding driving force Eq.2.15 and Eq.2.16. Out of the total maximum adsorption capacity q_{eq} of the sorbent, only $\psi\, q_{eq}$ is available on the surface of the sorbent particle. Accordingly, the driving force is modified as $\psi q_{eq} - q_{bulk}$. Similarly, the driving force for the amino-polymer sites is modified considering the remaining equilibrium adsorption capacity.

It is to be noted that the formulation of LDF model for the two adsorption sites as in Eq.2.15-Eq.2.16 is similar to the one reported in Bollini *et al* [52]. The difference is that, while Bollini *et al* [52] had used a lumped mass transfer coefficient parameter to describe the diffusion resistance, we have developed a more detailed mass transfer resistance model for each of the phase components and is described in the following section.

Overall mass transfer resistance

Under the assumption that the accumulation of $\mathrm{CO_2}$ within the macropores of the fiber is significantly smaller than the amount adsorbed in the sorbent, an overall mass transfer coefficient corresponding to the overall concentration driving force between the gas phase C_{g,co_2} and the sorbent phase q, is used in the present work. This, in effect, eliminates a partial differential equation that needs to be solved for the mass balance of the $\mathrm{CO_2}$ concentration in the fiber macropores. Fig.2.4b) schematically describes the individual phase resistances and the respective concentration gradients in each phase of the hollow fiber membrane. Despite of the simplification discussed above, the resistance offered by the fiber macropores, gas phase and the sorbent phase are rigorously accounted and the the overall mass transfer coefficient is obtained as sum of the individual resistances (as the

resistances are in series) for both the adsorption sites and are as follows.

$$\frac{1}{K_{ov,surf}} = \frac{1}{K_g} + \frac{1}{K_m} + \frac{1}{K_{s,surf}} \tag{2.19}$$

$$\frac{1}{K_{ov,bulk}} = \frac{1}{K_g} + \frac{1}{K_m} + \frac{1}{K_{s,surf}} + \frac{1}{K_{s,bulk}} \tag{2.20}$$

where $1/K_g$ and $1/K_m$ refer to the diffusion mass transfer resistances in the gas phase and the fiber macropore phase respectively. In addition, $1/K_{s,surf}$ and $1/K_{s,bulk}$ refer to the diffusion mass transfer resistances to the surface PEI layer and bulk PEI layer respectively. Fiber macropore resistance $1/K_m$ is henceforth referred to as inter-particle resistance, and the sorbent phase resistance is referred to as intra-particle resistances which are intraparticle surface diffusion resistance $1/K_{s,surf}$ and intra-particle polymer diffusion resistance $1/K_{s,bulk}$. The model for each of the component resistances is derived from the first principles and shown in Eq.(2.21-2.24)

$$K_g = \frac{2k_g\epsilon_f r_o}{(r_o^2 - r_i^2)\frac{\partial q_{eq}}{\partial C_{g,co_2}}\rho_f(1-\epsilon_f)} \tag{2.21}$$

$$K_m = \frac{8r_o^2 D_{fp}\epsilon_f}{(r_o^2 - r_i^2)^2\frac{\partial q_{eq}}{\partial C_{g,co_2}}(1-\epsilon_s)\rho_f(1-\epsilon_f)} \tag{2.22}$$

$$K_{s,surf} = \frac{15D_{p,k}\epsilon_s}{r_s^2\frac{\partial q_{eq}}{\partial C_{g,co_2}}\psi(1-\epsilon_s)\rho_s} \tag{2.23}$$

$$K_{s,bulk} = \frac{15D_p\epsilon_s}{r_s^2\frac{\partial q_{eq}}{\partial C_{g,co_2}}(1-\psi)\rho_s} \tag{2.24}$$

wherein k_g is the gas convective mass transfer coefficient, D_{fp} is the macropore gas diffusivity, $D_{p,k}$ is the Knudsen diffusion in the sorbent micropore and D_p is the sorbent polymer diffusivity. Knudsen diffusivity within the sorbent micropore is calculated as follows.

$$D_{p,k} = \frac{d_{s,pore}}{3}\sqrt{\frac{8R_gT_g}{\pi M_{co_2}}} \tag{2.25}$$

where $d_{s,pore}$ is the sorbent pore diameter and M_{co_2} is the molecular weight of CO_2. The gas convective mass transfer coefficient k_g is estimated within the model using the correlation relating Sherwood number $Sh = k_g r_h / D_g$ to Reynolds $Re = 2\rho_g u_g r_{fs}/\mu_g$ and Schmidt number $Sc = \mu_g/\rho_g D_g$ from Asimakopolou *et al*[56] as follows.

$$Sh = 1.45 \left(Re\, Sc\, \frac{d_o}{L} \right)^{0.33} \tag{2.26}$$

In the Sherwood number Sh, D_g refers to the molecular diffusivity of CO_2 in N_2 and calculated using the Chapman and Enskog equation [51]. Effective hindered molecular diffusivity in fiber macropore, D_{fp} is calculated using using the following equation.

$$D_{fp} = D_g \frac{\epsilon_f}{\tau_f} \tag{2.27}$$

Model for amino-polymer diffusion coefficient

Diffusion in the bulk PEI is considered slow, which causes a long tail in the CO_2 adsorption breakthrough curve. Unless the diffusion is modeled carefully, the model would not be able to predict the breakthrough curves. In this study, we apply a diffusion model to diffusivity in the amino-polymer phase of PEI, which is a long chain of amines.

Diffusivity in polymers have been most commonly estimated using free volume theory which considers the diffusivity to be proportional to the free volume available within the polymer as follows [57].

$$D_{polymer} \propto e^{-B/v_f} \tag{2.28}$$

where v_f is the free volume within the polymer molecule and the B is the volume of the penetrant molecule. Since poly(ethyleneimine) is a polymer that consists of a repeating unit of amine group and two aliphatic groups, we use an analogous formulation to model

the sorbent polymer diffusivity as follows.

$$D_p = D_{p0}^{'} e^{-\alpha q_{bulk}} \tag{2.29}$$

In the above equation, CO_2 concentration in the bulk PEI q_{bulk} can be understood as inversely proportional to free volume within the bulk PEI. Also, α can be understood as a parameter describing the rate at which sorbent CO_2 concentration reduces the free volume in the amino-polymer within the sorbent and thus effectively decreasing the polymer diffusivity with the progress of adsorption. Additionally, temperature has a significant influence on the diffusion coefficients $D_{p0}^{'}$ in polymer and hence the temperature dependency is also included according to Arrhenius relation as follows.

$$D_{p0}^{'} = D_{p0} e^{-E/R\,T_g} \tag{2.30}$$

where E is the activation energy for diffusion and R is the gas constant. Combining the above two equations, the sorbent polymer diffusivity D_p can be written as follows.

$$D_p = D_{p0} e^{-E/R\,T_g} e^{-\alpha q_{bulk}} \tag{2.31}$$

2.2.5 CO_2 adsorption isotherm

Equilibrium adsorption capacity of CO_2 in the sorbent phase, q_{eq} is determined using the Toth adsorption isotherm as follows.

$$\frac{q_{eq}}{q_m} = \frac{bC_{g,co_2}}{(1 + (bC_{g,co_2})^n)^{1/n}} \tag{2.32}$$

The temperature dependency of the affinity constant b, maximum sorption capacity q_m and heterogeneity constant n are modeled in the following equations,

$$b = b_0 e^{-\frac{\Delta H_0}{R_g T_0}\left(\frac{T_0}{T}-1\right)} \tag{2.33}$$

$$q_m = q_{m_0} e^{\left(\eta(1-\frac{T}{T_0})\right)} \tag{2.34}$$

$$n = A + B\left(1 - \frac{T_0}{T}\right) \tag{2.35}$$

where b_0 is the affinity constant at reference temperature T_0, ΔH_0 is the isosteric heat of adsorption at zero loading, η is the parameter defining the temperature dependency of maximum saturation capacity q_m, and q_{m0} is the maximum saturation capacity at T_0. Finally, A and B are the parameters defining the temperature dependency of the heterogeneity constant n.

2.3 Model implementation

The initial model development for the validation of the model to check for its adequacy was performed in gPROMS [58], which is a commercial dynamic process modeling and optimization software. The parameter estimation using the conventional method of Maximum Likelihood Estimation (MLE) was also performed in gPROMS. In gPROMS, the governing equations are solved using method of lines with a finite difference based discretization for the spatial derivatives. For all the simulations reported in this work, Backward Finite Difference (BFD) was used to discretize axial derivatives in all the equations except momentum balance and the Central Finite Difference (CFD) was used for the radial derivatives. The momentum balance equation is discretized using Central Finite Difference (CFD). However, as the UQ analysis requires a large number of simulations to be performed in parallel and the gPROMS licenses are limited, the entire RTSA cyclic process model along with the UQ algorithms are developed and implemented in Python using non-proprietary codes. In Python, the discretization schemes are explicitly coded and the basic accuracy of the

Table 2.1: Values of physical properties and process parameters

Physical Properties and parameters	**Symbol**	**Values**
Fiber heat capacity[J/kgK]	C_{pf}	1200
Flue gas heat capacity[J/KgK]	C_{pg}	1094
Flue gas viscosity [kg/ms]	μ_g	1.98×10^{-5}
Fiber bulk density with C803 [kg/m^3]	ρ_f	960
Fiber bulk density with ES757 [kg/m^3]	ρ_f	700
Fiber conductivity[W/mK]	k	0.05
Fiber tortuosity	τ_f	2.5
Fiber porosity	ϵ_f	0.48
Sorbent porosity (C803)	${\epsilon_s}^{c803}$	0.3
Sorbent density (C803) [kg/m^3]	${\rho_s}^{c803}$	800
Sorbent porosity (ES757)	${\epsilon_s}^{es757}$	0.08
Sorbent density (ES757)[kg/m^3]	${\rho_s}^{es757}$	200
Sorbent pore diameter[nm]	$d_{s,pore}$	20
Sorbent particle diameter(C803) [μm]	d_s	4
Sorbent particle diameter(ES757) [μm]	d_s	25
Volume loading (C803)(vol$_{sorbent}$/vol$_{fiber\ solids}$)	v_s	0.684
Volume loading (ES757)(vol$_{sorbent}$/vol$_{fiber\ solids}$)	v_s	2.2
Gas constant[J/mol K]	R_g	8.314
Thermocouple heat capacity[J/kg K]	C_{pt}	500
Thermocouple tip diameter[m]	d_t	0.9×10^{-3}
Thermocouple needle length[m]	L_t	0.05
Thermocouple conductivity [W/mK]	k_t	16
Average heat of adsorption [J/mol]	ΔH_{ads}	-59.0 $\times$ 10^3

implemented code in Python is verified by validating the results from Python code against the results from gPROMS model.

2.4 Cyclic simulation results

The values of the physical model parameters are given in Table.2.1. The fiber dimensions for the simulations of the large scale fiber is different from those used in the experimental set-up and are : r_i = 160 μ m, r_o = 600 μm, r_{fs} = 774 μm, L = 3 m. As, these are the dimensions that are required to predict the process performance in the plant-scale. The unknown parameters, viz: adsorption isotherm parameters, mass transfer resistance parameters and the heat transfer parameters, which are estimated by fitting the model against the experimental data, are fixed at the estimated values while performing the below cyclic simulation.

Table 2.2: Boundary conditions of the cycle for the first two steps

Adsorption	**Self-sweeping**
$T_w\|_{z=0} = T_{ads}, \frac{\partial T_w}{\partial z}\|_{z=L} = 0$	$T_w\|_{z=0} = T_{des}, \frac{\partial T_w}{\partial z}\|_{z=L} = 0$
$T_g\|_{z=0} = T_{ads}, \frac{\partial T_g}{\partial z}\|_{z=L} = 0$	$T_g\|_{z=0} = T_{des}, \frac{\partial T_g}{\partial z}\|_{z=L} = 0$
$C_{g,co_2}\|_{z=0} = y_{co_2}\frac{P}{RT_g}\|_{z=0}, \frac{\partial C_{g,co_2}}{\partial z}\|_{z=L} = 0$	$\frac{\partial C_{g,co_2}}{\partial z}\|_{z=0} = 0, \frac{\partial C_{g,co_2}}{\partial z}\|_{z=L} = 0$
$C_{g,N_2}\|_{z=0} = y_{N_2}\frac{P}{RT_g}\|_{z=0}, \frac{\partial C_g}{\partial z}\|_{z=L} = 0$	$\frac{\partial C_{g,N_2}}{\partial z}\|_{z=0} = 0, \frac{\partial C_{g,N_2}}{\partial z}\|_{z=L} = 0$
$u_g\|_{z=0} = u_{ads}, \frac{\partial u_g}{\partial z}\|_{z=L} = 0$	$u_g\|_{z=0} = 0, \frac{\partial u_g}{\partial z}\|_{z=L} = 0$
$\frac{\partial P}{\partial z}\|_{z=0} = 0, P\|_{z=L} = 1.013 \times 10^5$	$\frac{\partial P}{\partial z}\|_{z=0} = 0, P\|_{z=L} = 1.013 \times 10^5$

Table 2.3: Boundary conditions of the cycle for the last two steps

N_2 **sweeping**	**Cooling**
$T_w\|_{z=0} = T_{des}, \frac{\partial T_w}{\partial z}\|_{z=L} = 0$	$T_w\|_{z=0} = T_{ads}, \frac{\partial T_w}{\partial z}\|_{z=L} = 0$
$T_g\|_{z=0} = T_{des}, \frac{\partial T_g}{\partial z}\|_{z=L} = 0$	$T_g\|_{z=0} = T_{ads}, \frac{\partial T_g}{\partial z}\|_{z=L} = 0$
$C_{g,co_2}\|_{z=0} = 0, \frac{\partial C_g}{\partial z}\|_{z=L} = 0$	$C_g\|_{z=0} = 0, \frac{\partial C_{g,co_2}}{\partial z}\|_{z=L} = 0$
$C_{g,N_2}\|_{z=0} = 1.0\frac{P}{RT_g}\|_{z=0}, \frac{\partial C_{g,N_2}}{\partial z}\|_{z=L} = 0$	$C_{g,N_2}\|_{z=0} = 1.0\frac{P}{RT_g}\|_{z=0}, \frac{\partial C_{g,N_2}}{\partial z}\|_{z=L} = 0$
$u_g\|_{z=0} = u_{purge}, \frac{\partial u_g}{\partial z}\|_{z=L} = 0$	$u_g\|_{z=0} = u_{cool}, \frac{\partial u_g}{\partial z}\|_{z=L} = 0$
$\frac{\partial P}{\partial z}\|_{z=0} = 0, P\|_{z=L} = 1.013 \times 10^5$	$\frac{\partial P}{\partial z}\|_{z=0} = 0, P\|_{z=L} = 1.013 \times 10^5$

The parameter estimation results and the values thereby obtained are discussed in the next chapter and are not repeated here. The thermal swing is carried out by sending cold and hot water through the fiber bore during the adsorption and desorption steps respectively. The temperature is swung between 55^oC during adsorption and 140 ^oC during desorption. The boundary conditions that describe the process operation during each of the steps in the cycle for an industrial-scale of the process are described in Tables 2.2 and 2.3. The operating process conditions for the various process variables are listed in Table 2.4.

The simulation is carried out until the cyclic steady state (CSS) is reached, by which, the values of the state variables at the start of the cycle and the end of the cycle need to be the same. The step times are tuned so as to obtain maximum purity and recovery

Table 2.4: Operating conditions of the process

Process variable	**Value**
u_{ads}	4.5 m/s
$u_{N-sweep}$	0.045 m/s
u_{cool}	1.2 m/s
$u_{w,ads}$	0.5 m/s
$u_{w,s-sweep}$	0.5 m/s
$u_{w,N-sweep}$	0.3 m/s
$u_{w,cool}$	0.5 m/s

possible and are as follows: adsorption is carried out for 50 seconds, followed by self sweeping for 62 seconds, followed by N_2 sweeping for 33 seconds and finally cooling for 50 seconds, which sums up to the RTSA cycle time of 195 seconds. One of the observable from the experimental set-up is the CO_2 breakthrough molar percentage at the exit of the fiber. Figure 2.5 shows the CO_2 molar percentage at the exit of the fiber over the cycle. The outlet CO_2 is recovered during most of the self-sweeping step and throughout the N_2 sweeping step (marked by green solid lines in the figure), resulting in about 64% recovery of the total inlet CO_2 with a purity of 96%.

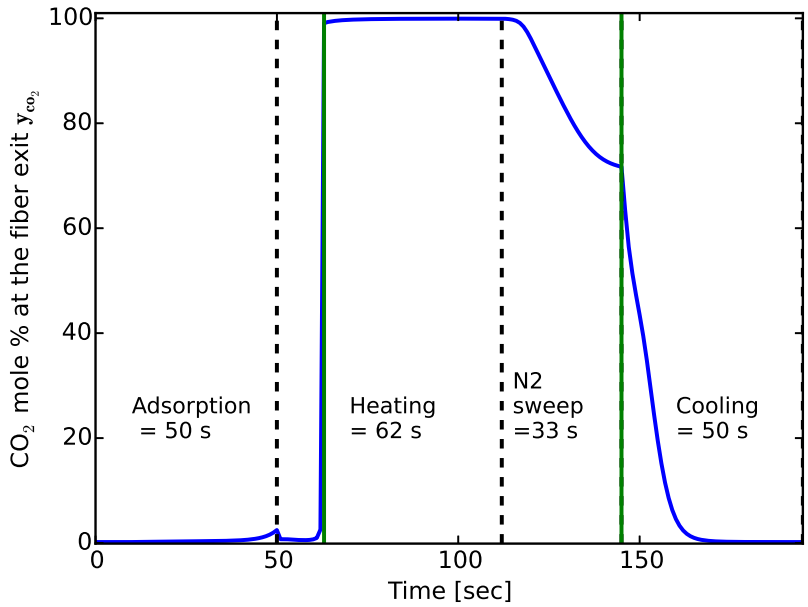


Figure 2.5: Molar percentage of CO_2 at the fiber exit over the cycle. CO_2 is recovered from the outlet stream in the duration indicated by green solid lines.

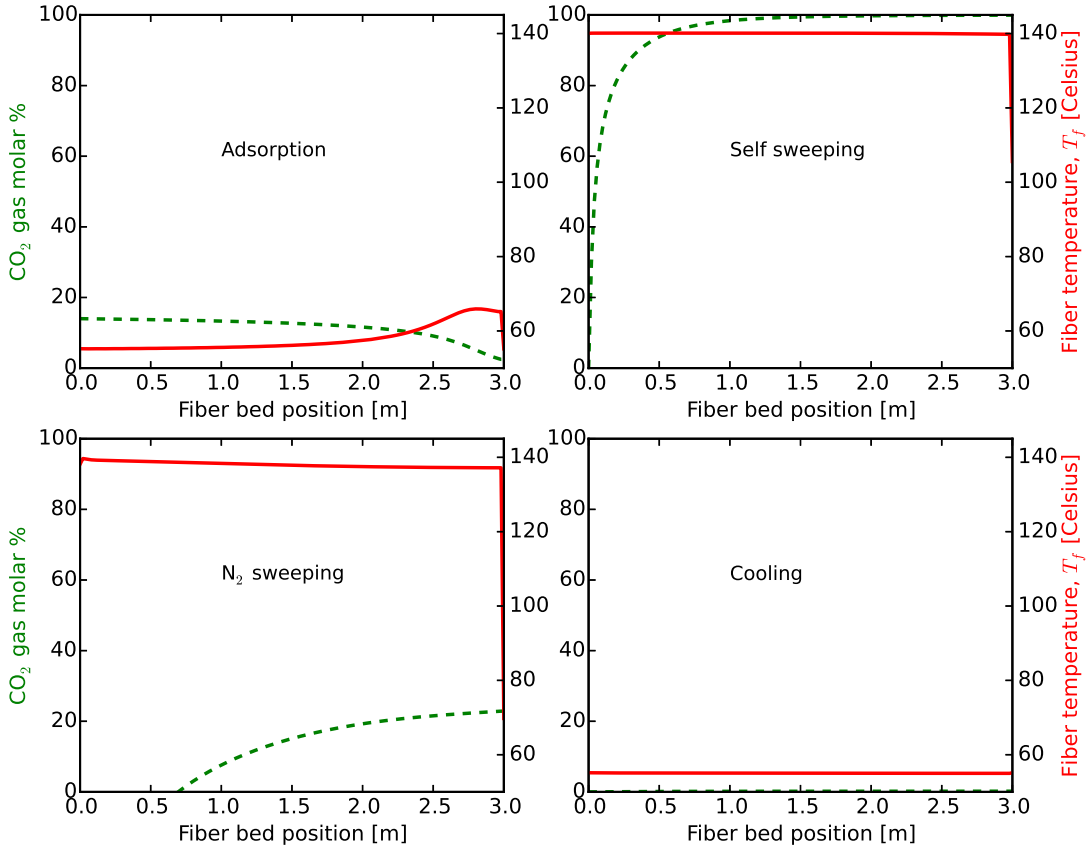
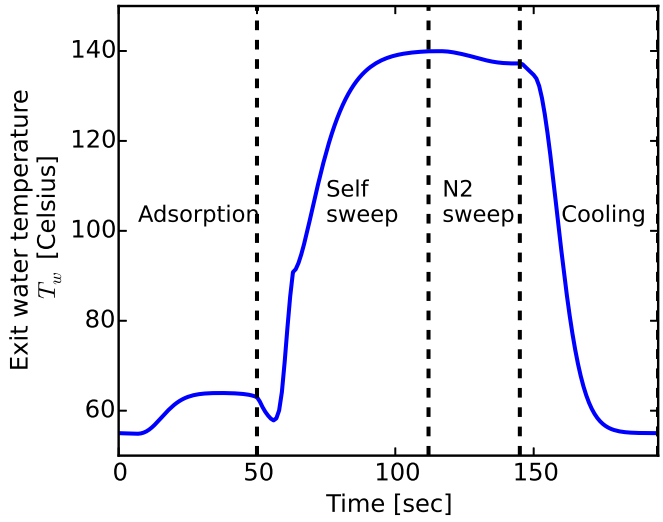


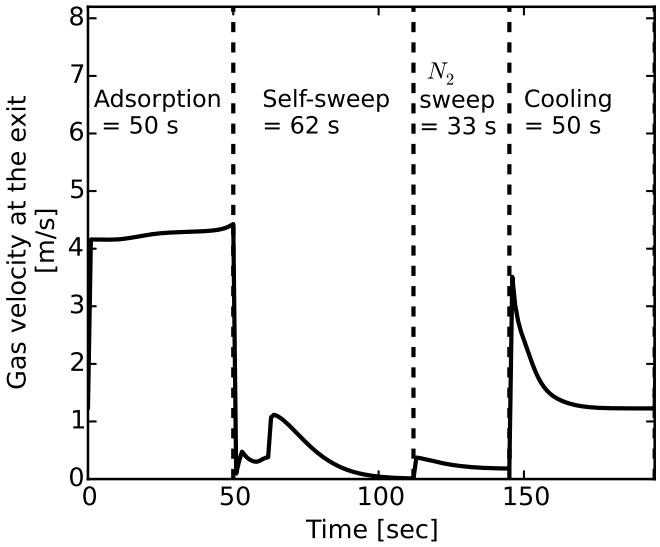
Figure 2.6: CO_2 concentration and fiber temperature variations along the bed position at the end of each step.

Simulations can also be used to visualize other physical variables in the model, those that cannot be easily measured in the experiments, however can provide more insight into the physics of the process. Figure 2.6 shows the variation of gas phase CO_2 concentration and the fiber temperature along the bed position at the end of each of the steps in each sub-figure. The water flow rate is maintained at a high value for both adsorption and desorption steps, in order to maximize the CO_2 swing capacity which is the working CO_2 capacity. As a result, the observed rise in fiber temperature from the release of heat of adsorption, is not significant. Similarly, there is no significant drop in temperature during the desorption steps from the consumption of heat of desorption. Figure 2.7(a) shows the variation of water temperature at the exit of the fiber during the cycle, which could be measured in the experiment. However, as the experiments are all conducted without the flow of water in the

bore, this data is not available.



(a)



(b)

Figure 2.7: a) Water temperature at the fiber exit during the cycle, b) Gas velocity at the fiber exit during the cycle.

The gas velocity at the fiber exit is shown in Figure 2.7(b), which indicates the flow rate during the adsorption steps are much higher than the desorption steps. The initial rise in the self-sweeping step is due to the bulk flow of released CO_2, which then decreases due to the lack of sufficient pressure drop. The spike in the beginning of the cooling step is due to the

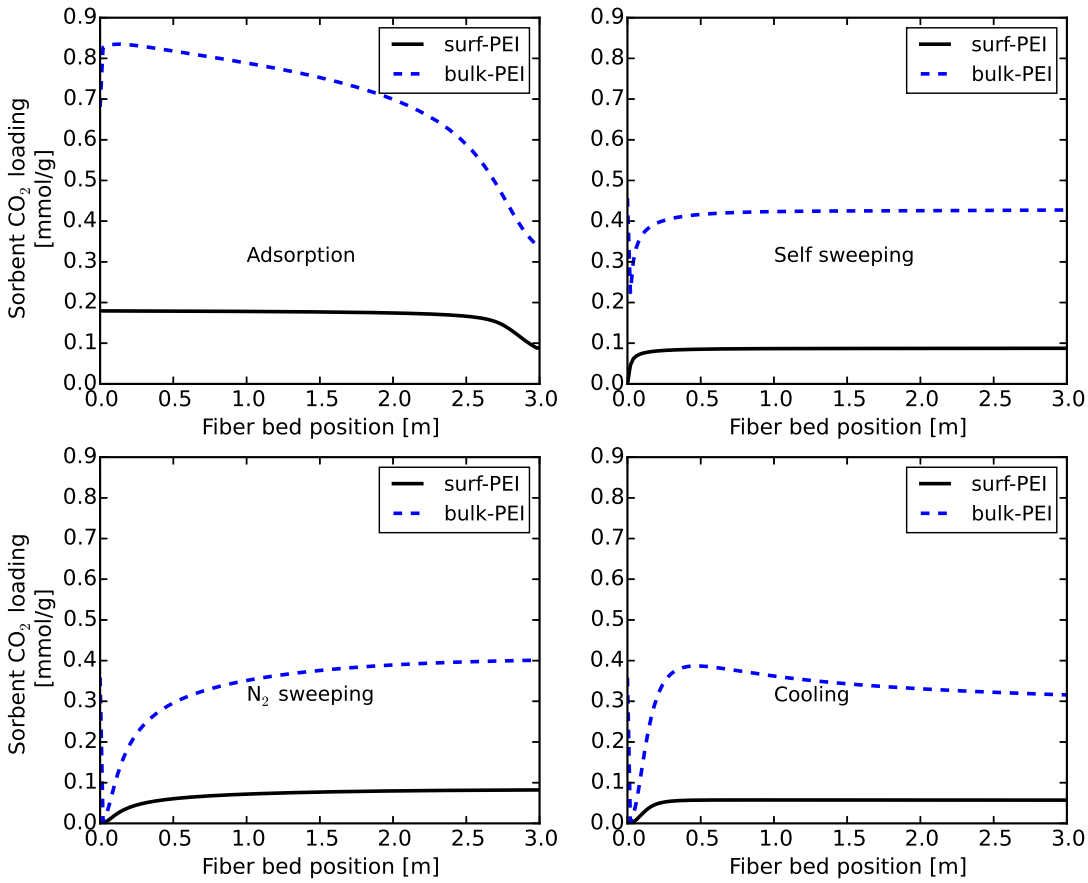


Figure 2.8: CO_2 loading in the sorbent in the two different amine sites at the end of each step.

transient buildup of mass and pressure due to the sudden increase of inlet flow rate while transitioning from the N_2 sweep to the cooling step. Based on the mass transfer model, which assumes that the amine sites are available as two different sites offering different scales of diffusion resistances, the sorbent loading profiles of the two different sites are obtained and are shown in Figure 2.8. As can be seen from the figure, the surface sites are of a small fraction and have a low mass transfer resistance and therefore are more sharp in their fronts. On the other hand, the bulk amine loading is a larger fraction and has a diffusive profile due to the higher diffusion mass transfer resistance. The other key performance metric that is of significance is the swing (or working) capacity of adsorbent, which is the effective adsorption capcity of sorbents that is being utilized between adsorption and desorption. The working capacity or the swing capacity, q_{swing,co_2} that is obtained with this cycle conditions is 0.62 mmol/gfib.

The process performance predictions obtained, thus far, based on the point estimate values of physical model parameters, however, may not be completely reliable given that there are uncertainties both in experimental data and in model, whose effect are not incorporated in the model prediction. The next chapter describes the conventional methods used to determine the point estimates of model parameters and why such point estimates based model may not be sufficient in applications such as scale-up studies, techno-economic feasibility etc.

CHAPTER 3

CONVENTIONAL PARAMETER ESTIMATION AND ITS LIMITATIONS

The physical parameters of the RTSA model Θ, whose values are unknown, are estimated by fitting the model against the experimental data. The parameter estimation is initially done using conventional techniques, viz. Maximum Likelihood Estimation (MLE) [59] and Ordinary Least Squares (OLS)[60] method, which has helped to serve two purposes. Primarily, it provides a good prior information on the range of values that have to be explored while performing the parametric inference under UQ analysis. In addition, it also helps to understand if the model is structurally adequate (or adequately parametrized) to explain the experimental data.

3.1 Ordinary Least Squares Regression (OLS)

The most commonly used and well known method for estimating parameters in classical statistics is by minimizing the sum of squares of the differences between the observed responses, y, and the corresponding model predictions, $\hat{y}$ based on a linear model $y = X\beta_p + \epsilon$ with a set of predictor or dependent variables, x_i. Here, X is a matrix of size $m \times n$, whose m rows correspond to m dependent variables, $x_i, i = 1..m$, and n columns to the values of the m dependent variables corresponding to each observations of $y_j, j = 1..n$. The linearity assumption is only with respect to the parameters and the dependent or the predictor variables can be highly non-linear as long as their parametric coefficients are linear with one another. The other key assumption in OLS method is that model mismatch

errors ϵ are normally distributed with zero mean and variance σ_y^2 as $\mathcal{N}(0, \sigma_y^2)$. However, in case of data having non-Gaussian distributions, such as binomial or Poisson, Generalize Least Squares (GLM) method is used to estimate the parameters [60].

The parameters β_p are estimated using matrix algebra as $\hat{\beta}_p = (X^T X)^{-1} X^T y$. Under the assumption of normality of data y, the parameter estimates also follow Gaussian distribution as $\mathcal{N}(\beta_p, \sigma_y^2 (X^T X)^{-1})$ indicating the mean of parametric estimate is its true value, β_p and the variance is proportional to the variance of data, y. As the parameters are obtained as "point estimates" $\hat{\beta}_p$ from OLS, confidence intervals are constructed to determine the uncertainty around the estimates. Typically, as the standard error value σ_y is unknown, t statistic is used to calculate the confidence intervals as $\hat{\beta}_p \pm t_{1-\frac{\alpha}{2}} \hat{\sigma}_y \sqrt{(X^T X)^{-1}}$, where t is calculated at *df* = $n - m - 1$ and with $\alpha = 0.05$ for 95 % confidence.

3.2 Maximum Likelihood Estimation (MLE)

In case of models with non-linear parametric dependence, OLS method is not applicable. In such cases, Maximum Likelihood Estimation (MLE) with Gaussian distribution [59] is most commonly used to estimate parameters, which actually reduces to OLS estimates, when the ϵ in OLS are assumed to be normally distributed. Under MLE, the parameters that have a higher likelihood (or probability) to fit the model are estimated by maximizing the likelihood function; $L(y|\theta)$, a joint probability distribution of data y being predicted by the model that is defined by parameters θ. Gaussian likelihood is accordingly defined as a multivariate probability distribution of observed data, y with mean at its model prediction value y_m and variance σ_y^2 and is described as follows.

$$\mathcal{L}(\theta|\mathrm{y}) = \frac{1}{\sqrt{2\pi\sigma_{\mathrm{y}}^2}^{\mathrm{n_{exp}}}} \exp\left(-\sum_{\mathrm{i}=1}^{\mathrm{n_{exp}}} \frac{(\mathrm{y_i} - \mathrm{y_{m,i}}(\theta))^2}{2\sigma_{\mathrm{y}}^2} \right) \tag{3.1}$$

The underlying assumption in the above likelihood is that the data y are uncorrelated and have equal variance σ_y^2 (homoscedasticity property).

The function is maximized and can be solved using algorithms such as Marquardt-Levenberg algorithm or Gauss Newton to determine the optimum. Again, the parameters are obtained as point estimates and the uncertainty estimate is obtained by constructing the 95% confidence interval using the t statistic as explained before.

3.3 Estimation of model parameters

There are welve physical model parameters to be estimated: six adsorption equilibrium parameters θ_{iso}, four mass transfer parameters θ_{mass} and three heat transfer parameters θ_{heat}. It is quite challenging to estimate all the unknown parameters simultaneously as there are a large number of parameters to be estimated. Besides, there also exists strong correlation among some of the parameters which can make the estimation more difficult. Therefore, the parameters are estimated by conditionally isolating some of them and designing the experiments accordingly to gather the required data.

3.3.1 Estimation of adsorption isotherm parameters

The six adsorption isotherm parameters can be estimated separately by conditionally isolating the adsorption equilibrium model given in Eqns. 2.32-2.35, at various temperatures and CO_2 partial pressures. As CO_2 adsorption equilibrium capacity can be directly measured from Thermo Gravimetric Analysis (TGA) experiments at various temperatures and CO_2 partial pressures, those data are obtained and are used to estimate the adsorption isotherm parameters θ_{iso}. The adsorption isotherm parameters in Eqns. 2.32- 2.35, which are estimated using the OLS method with the TGA data are $[q_{m0}, \eta, A, B, b_0, \Delta H_0]$. The estimation is done using the inbuilt Python function *curvefit* of *scipy.optimize* package [61], which uses Levenberg-Marquardt algorithm to minimize the least squares of errors. The model fit obtained with the experimental data during the estimation is shown in Fig. 3.1. The estimated values of the adsorption isotherm parameters are given in Table. 3.1 along with the 95% Confidence Intervals (CI) of the OLS estimates. These confidence intervals

Table 3.1: Adsorption isotherm parameters estimated by OLS method

Method	Value	95% CI [LB, UB]
$q_{m0}[mmol/gfib]$	1.330	[1.230, 1.430]
η	1.930	[1.270, 2.590]
A	0.506	[0.387, 0.625]
B	3.430	[1.552, 5.310]
$b_0 \times 10^3[\text{bar}^{-1}]$	1.109	[0.151, 2.069]
$\Delta H_0[kJ/mol]$	-65.6	[-77.2 , -54.1]

are obtained based on the assumption in the OLS estimation that the errors are normally distributed. Accordingly, the parameters are also normally distributed with mean as the estimate values and variance given by $\sigma_{q_{eq}}\sqrt{(X'X)^{-1}}$. Here, X refers to the values of the independent (or the predictor) variables in the isotherm model, which are the CO_2 partial pressures and temperatures.

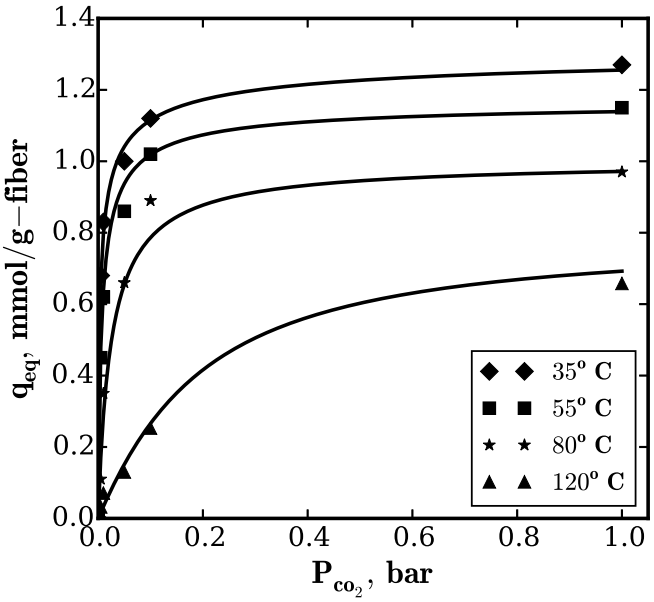


Figure 3.1: Model prediction with the OLS estimates of adsorption isotherm parameters

3.3.2 Estimation of mass and heat transfer parameters

As the mass transfer parameters which need to be estimated $\theta_{mass} = [D_{p0}, E, \psi, \alpha]$ are highly inter-correlated, it is difficult to estimate all of them simultaneously. Therefore,

some of the parameters are isolated and estimated separately.

Estimation of ψ, fraction of surface amine sites

The parameter ψ is estimated independently by using some existing measurements from literature of related physical properties. It denotes the fraction of the amine sites that are available closer to the void space of the pores and therefore easily accessible for CO_2 adsorption. As the available experimental data in literature on ψ [55] is for PEI impregnated in a different silica (SBA-15), reverse engineering calculations were made to determine the value of ψ for PEI impregnated C803 silica (which is used in this thesis) based on the measured ψ of SBA-15. The amine loading values and the surface area measurement of C803 that are used in the calculation are obtained from Labreche *et al* [11]. Based on the available measurements, value of ψ is calculated to be 0.17, i.e 17% of the total amine (PEI) sites are available on the surface of the C803 silica pores. The details of the calculation can be found in the Appendix A.

Estimation of sorbent-polymer diffusivity and heat loss parameters

The sorbent polymer diffusion parameters D_{p0}, E and α along with the heat loss parameters h_t, U and axial dispersion coefficient D_{ax} are estimated by matching the CO_2 breakthrough and fiber temperature profile measurements with their respective model prediction. Among these parameters, the temperature dependency parameter E requires the estimation to be conducted with CO_2 breakthrough at several temperatures. However, it is computationally very expensive to perform the estimation with several CO_2 breakthrough data simultaneously due to the larger size of the model involved. Besides, the correlation between the mass transfer parameters also poses challenge in estimating all of them simultaneously. Therefore, the estimation of the parameters are carried out in sequence.

Initially, the CO_2 breakthrough data and fiber temperature profiles measured at a flue gas flow rate of 120 mL/min and 35 ^{o}C using a hollow fiber module of length $L = 10.25\,in$

Table 3.2: List of Mean Square Error (MSE) values for model fit

Experiment	Conditions	MSE
Exp1	10-inch, 120 mL/min, 35^{o} C	2.5
Exp2	10-inch, 120 mL/min, 45^{o} C	0.03
Exp3	10-inch, 120 mL/min, 55^{o} C	0.04
Exp4	10-inch, 120 mL/min, 65^{o} C	0.09

is used to estimate the parameters D_p (and not D_{p0}), α, h_t, and U alone leaving out E. Later, the parameter estimation is repeated with CO_2 breakthrough and fiber temperature data measured with the same module and at the same flow rate of 120 mL/min, however at different module temperatures viz: 45 ^{o}C, 55 ^{o}C and 65 ^{o}C individually. The latter parameter estimations that were carried out with breakthrough data at different module temperatures are done to estimate only D_p while the values of other parameters viz: α, h_t, U, and D_{ax} are retained at their previously estimated values. The underlying assumption, here, is that only D_p varies significantly with variation of module temperature, while the other parameters are relatively insensitive to the module temperature variation. It is to be noted that although each experiment, by itself, is non-isothermal, the overall temperature variations were minimal due to the large amount of heat losses to the module. The MLE in gPROMS is performed by minimizing the errors between the experimental CO_2 breakthrough and fiber temperature data against the respective predictions from the entire hollow fiber model Eq. 2.1-Eq. 2.35, which is run for adsorption conditions without the flow of cooling water. Figures 3.2-3.3 show the model fit obtained with CO_2 breakthrough and fiber temperature data at 35 ^{o}C, 45 ^{o}C, 55 ^{o}C and 65 ^{o}C and the respective mean square error between the model fit and the experimental data is listed in Table. 3.2.

With the estimated values of D_p at various module temperatures, the parameter E is then estimated by isolating the equation, Eq.2.30 and performing the OLS regression with that model. Figure 3.4 shows the model fit and prediction obtained to estimate the parameter E using OLS regression on the model Eq.2.30. Table 3.3 lists all the mass and heat transfer parameters, thus far estimated along with their 95% confidence intervals.

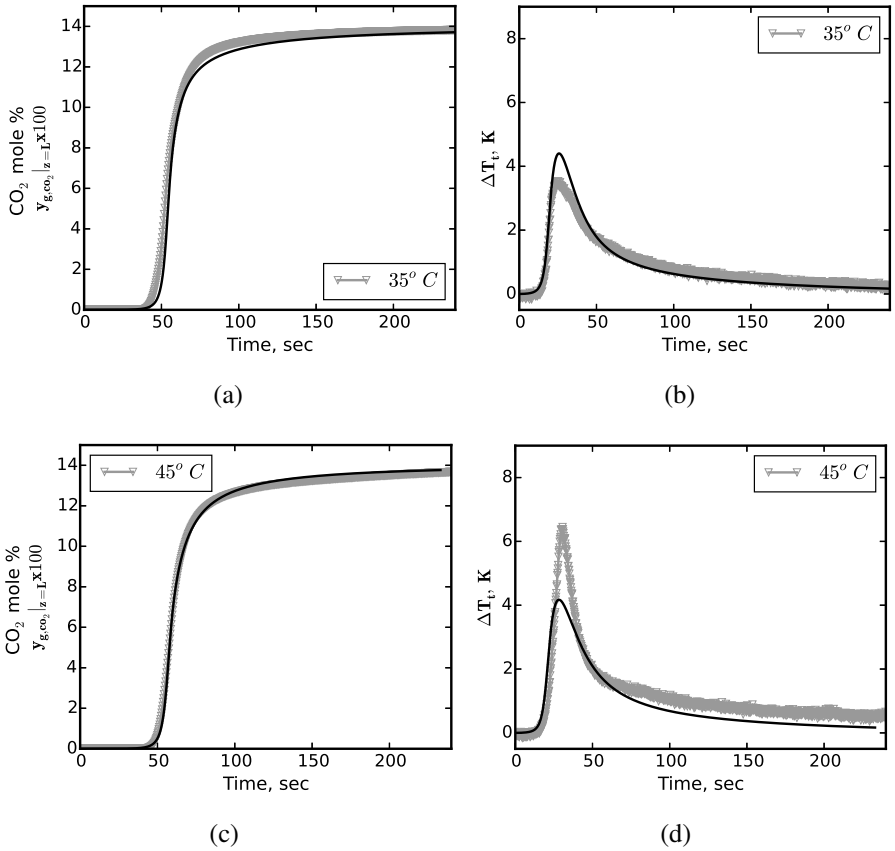


Figure 3.2: Parameter estimation via model fitting of a) CO_2 breakthrough at 35^oC b) T_f at 35^oC c) CO_2 breakthrough at 45^oC d) T_f at 45^oC

Table 3.3: Estimated mass and heat transfer parameter values

Parameters	Values	95% CI [LB, UB]
D_{p0} $[m^2/s]$	32.70	[0.53, 2025]
E $[kJ/mol]$	6.14	[4 .80, 7.50]
α $[gfib/mmol]$	7.13	[6.94, 7.28]
U $[W/m^2K]$	40.0	[30, 50]
$h_t[W/m^2K]$	1310	[1260, 1360]

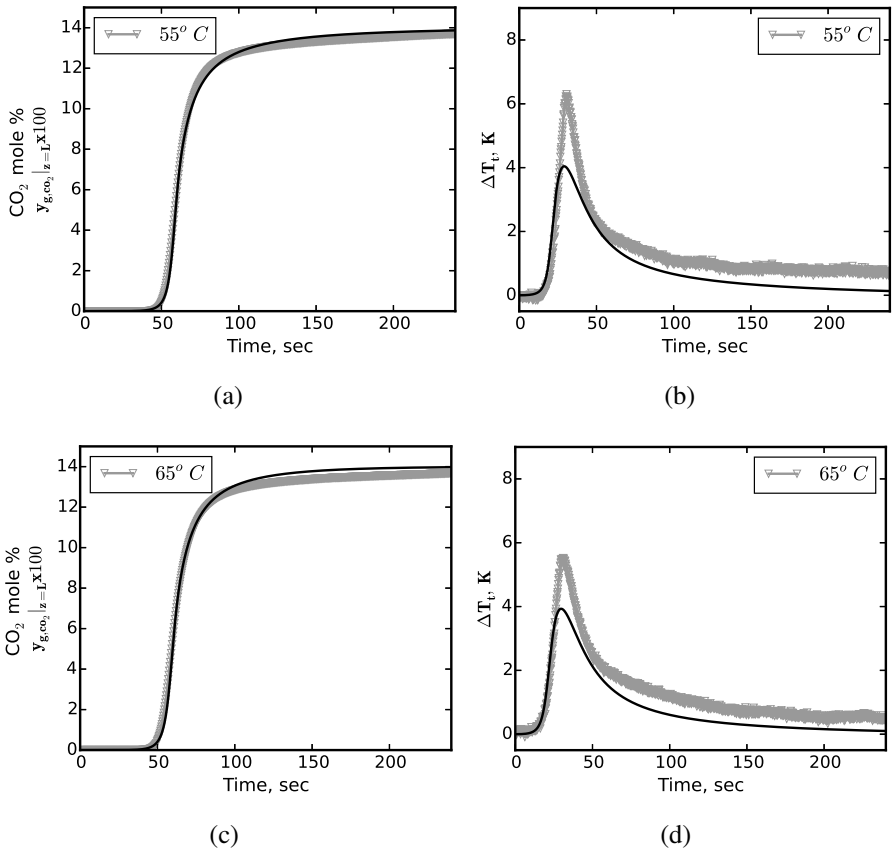


Figure 3.3: Parameter estimation via model fitting of a) CO_2 breakthrough at 55^oC b) T_f at 55^oC c) CO_2 breakthrough at 65^oC d) T_f at 65^oC

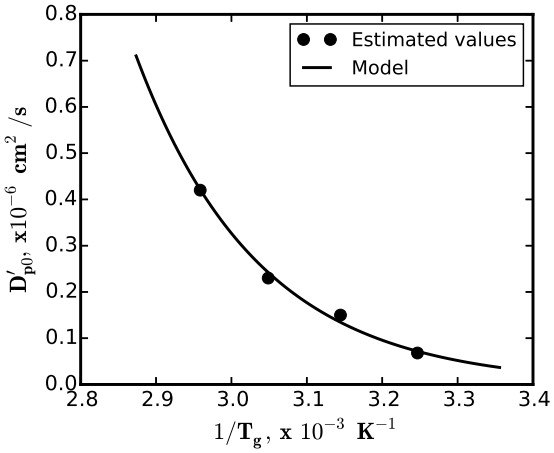


Figure 3.4: Model fit to estimate parameter E by OLS regression of Eq.2.30

3.4 Experimental model validation

The estimated model parameters are validated using new sets of experimental data collected with a different module of length 17 inch and housed with the same number of fibers and dimensions as the 10 inch module used for estimation. In addition to validating with the new data, the validation was also performed against the 10 inch module data, which are the repeated measurements under the same conditions as used for parameter estimation. A quantitative assessment of the goodness of model match is provided by calcuating the mean square error (MSE) between the model prediction and the experimental data and are listed for all the experimental data in Table. 3.4. As seen from Figs: 3.5 and the corresponding values in Table. 3.4, while the model matches very well for some of the experimental conditions (e.g: 10 inch module at $55^{o}\mathrm{C}$ and $65^{o}\mathrm{C}$ and 17 inch module at 200 mL/min), the model prediction is not very good at other experimental conditions.

The model validation was conducted by matching only the CO_2 breakthrough measurements at various conditions against the model predictions. The fiber temperature measurements, on the other hand, were ignored in the validation as they were found to involve huge uncertainty in the placement of thermocouple needle tip inside the module, affecting the thermal peak significantly. Nevertheless, the fiber temperature measurements were still used along with CO_2 breakthrough data to estimate the mass and heat transfer parameters. As can be seen from Fig. 3.6, the uncertainty in thermocouple position causing variation in the heat transfer coefficient value, h_t between the thermocouple and fiber does not have any influence on the position of thermal peak, but only varies the magnitude of the peak. Moreover, it is the position of thermal peak which determines the thermal front velocity rather than peak magnitude. Also, given that the thermal front velocity relative to mass transfer front is more sensitive while determining the parameter estimates, the uncertainty in fiber temperature magnitude is not expected to cause any bias on the estimated values.

Table 3.4: List of Mean Square Error (MSE) values for model predictions

Experiment	Conditions	MSE
Exp1	10-inch, 120 mL/min, 35° C	305.7
Exp2	10-inch, 120 mL/min, 45° C	209.5
Exp3	10-inch, 120 mL/min, 55° C	82.09
Exp4	10-inch, 120 mL/min, 65° C	57.38
Exp5	17-inch, 80 mL/min, 35° C	200.8
Exp6	17-inch, 120 mL/min, 35° C	234.2
Exp7	17-inch, 160 mL/min, 35° C	225.7
Exp8	17-inch, 200 mL/min, 35° C	179.6

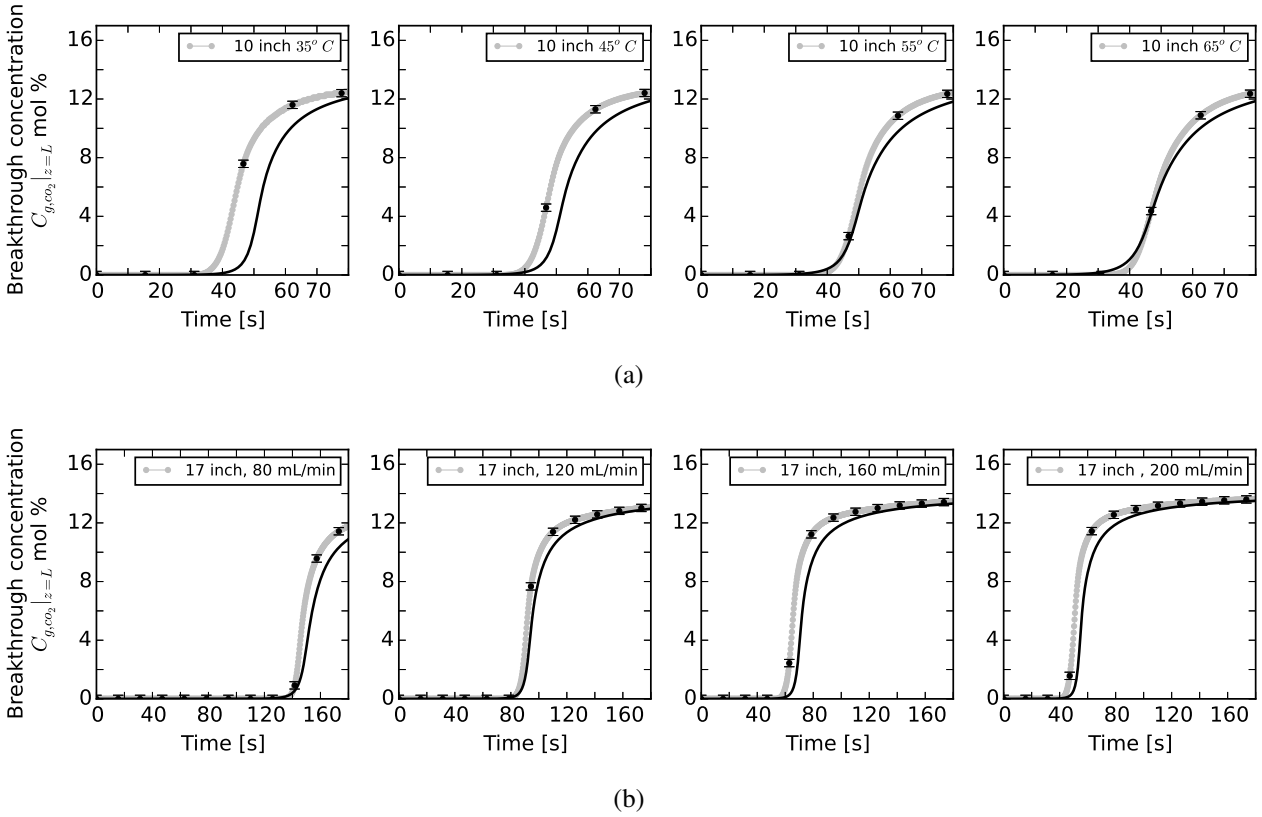
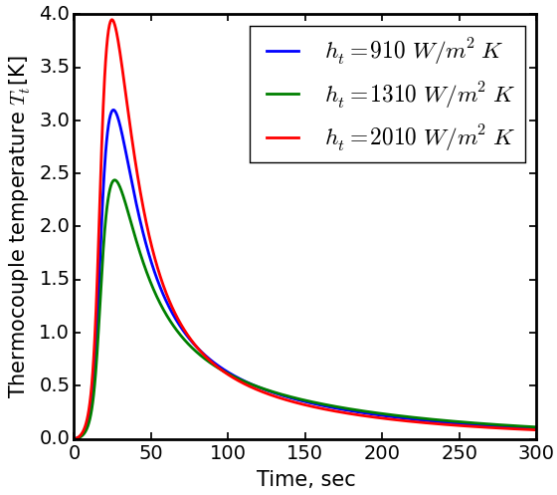


Figure 3.5: Experimental validation of estimated model parameters against new sets of experimental data a) from 10 inch module at various ambient temperatures b) from 17 inch module at various flue gas flow rates. Solid lines are the mean values of model predictions. Error bars are shown only on some data points as a representation.



(a)

Figure 3.6: Effect of thermocouple tip uncertainty causing variation in thermal peak magnitude.

3.5 Limitations of conventional parameter estimation

As shown under model validation section, the physical model parameters estimated using conventional methods of classical statistics viz: MLE and OLS, could not explain successfully all the experimental data that are measured at various operating conditions. It is also to be noted that the estimation could be conducted with only few experimental runs, the CO_2 breakthrough curves, as the computation increases with addition of experimental data. For processes requiring tremendous effort and resources to conduct experiments, the data may be gathered sequentially as needed. In such cases, there exists no methodology or provision in the conventional methods to refine the old parameter estimates by including only the newly gathered data. In order for the estimates to be consistent across all the experimental data used, all of them have to be included in the estimation simultaneously. Also, there exists no way to include the prior knowledge on the parameters, besides giving an initial guess or bounds on the parameter estimates.

With regard to model fitting, the predictions did not match the experimental data at conditions which are different than those used in estimation as can be seen in Fig.3.5(b),

although the model fitted very well with the experimental data during estimation having lowest MSE values (Table.3.2) compared to those obtained during validation (Table.3.4). Besides, the experiment also lacks reproducibility, due to which there is a model mismatch with some of the repeat measurements of the data that were used in parameter estimation as seen in Fig.3.5(a). Lack of reproducibility in data is due to the complexity of the process, as it involves several manual steps right from the synthesis of sorbent to be loaded in the hollow fibers, to the fabrication of hollow fibers with sorbents impregnated and finally in the housing of fibers within a module and location of the thermocouple. Owing to the presence of various errors and variability in the measurement data, it is impossible to explain all the experimental data based on 'point estimates' of model parameter values obtained as best estimates via MLE or OLS. Therefore, it becomes necessary to incorporate uncertainty in the estimates of parameter values such that all the data can be explained within the range of parameter values that are evaluated as likely estimates.

In conventional methods, the uncertainty in the estimates of parameters is provided via confidence intervals, which is the interval that contain the true value of the parameters with certain probability, on repeating the experiment for a large number of times. The calculation of confidence intervals is based on the assumption that the errors are normally distributed. As a result, parameters estimated also are normally distributed with their true value as mean and variance as $\sigma_y \sqrt{(X'X)^{-1}}$. This assumption of normal distribution for parameters, however, is valid only when the model involves linear parametric dependence. However, in most practical models, as in the case of carbon capture via RTSA process, the model involves nonlinear parametric dependence. In such cases, the uncertainty estimate provided by MLE or OLS is only approximate and cannot be used reliably as a measure to determine their impact on model prediction performance.

Therefore, a rigorous estimation method based on Bayesian principle is required to quantify the uncertainties in the parameter estimates and thereby in the model predictions. Besides, Bayesian inference also offers several other advantages such as sequential addi-

tion of experimental data in the parameter estimation and inclusion of prior information of parameters, both of which can be very helpful in cases of parametric inference for complex and large scale models. The next chapter describes the Bayesian inference and the different approaches that are used in the parametric calibration to incorporate the impact of uncertainties in the model, which arises due to various errors in the measurement and inadequacies in the model structure.

CHAPTER 4

UQ ANALYSIS: CHARACTERIZATION OF VARIOUS UNCERTAINTIES

As described in the previous chapter, the impact of various uncertainties must be incorporated in the model predictions, in order for the model to explain all the experimental data. On incorporating the effect of various uncertainties in the model, the predictions also become uncertain, thereby enabling the analysis of worst and best-case scenario or any other component of the distribution of outcomes. In effect, the reliability of model predictions is established via quantification of error margins and uncertainties (QMU), which is a very useful measure in decision making.

The most natural way to characterize uncertainties is by using probabilistic framework. This is because, most of the uncertainties can be attributed to random variability in the process, in which some of them can be controlled and reduced effectively (e.g: variability due to manual steps involved in synthesis of sorbent resulting in uncertainty in fiber adsorption capacity). Under the probabilistic framework, Bayesian approach provides a unified methodology to perform characterization, propagation as well as reduction of uncertainties.

4.1 Bayesian inference

The characterization of uncertainties, is the first step towards uncertainty quantification (UQ). As the values of physical model parameters are inferred while simultaneously characterizing the uncertainties associated with them, this step is also referred to as the parametric inference problem. The step involves determining the model input, i.e the model

parameters based on the model output, i.e. the experimental observation , thus it is classified as an inverse problem. The forward problem, on the other hand, involves determining the model output, which are the predictions of observations given some values of the physical parameters as model inputs.

Parametric inference via Bayesian approach involves determining the posterior distribution of parameters as follows (Gelman *et al.*. 2011):

$$P(\theta|y) = \frac{\mathcal{L}(y|\theta)P(\Theta))}{\int_\theta \mathcal{L}(y|\theta)P(\theta)d\theta}. \tag{4.1}$$

In the above equation $\mathcal{L}(y|\theta)$ is termed as the likelihood distribution, which is analogous to probability distribution in its formulation, where $\mathcal{L}(\theta|y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(\mathrm{y}-\mathrm{y_m}(\theta))^2}{2\sigma^2}\right)$. Here, the likelihood is the formula for a Gaussian probability distribution. However, the subtle difference between probability distribution and likelihood is that in probability the observation, y is a random variable and parameters, θ are fixed quantities. On the other hand, in likelihood, the observation y is fixed and likelihood indicates the probability of the model prediction, y_m with parameters θ to fit the data, y. The likelihood distribution is, therefore, evaluated by running the model simulation and calculating the error between the model prediction and experimental observation. The term $P(\theta)$ is the prior probability distribution of parameters, which includes the assumptions or prior knowledge of parameters before updating with the experimental data. The denominator is the integral of the numerator, marginalizing over all the parameters resulting in the total likelihood distribution of all the experimental data over all parameters. This term is also sometimes referred to as the evidence distribution. The resultant distribution from this evaluation on the left hand side of Eq. 4.1 is called the posterior probability distribution of parameters, as it is obtained by updating the prior distribution posterior to observing the experimental data.

4.2 Algorithms to perform parametric inference

The posterior distribution, $P(\theta|y)$, in case of non-linear complex models cannot be determined analytically. Therefore, numerical sampling methods such as Markov Chain Monte Carlo (MCMC), Sequential Monte Carlo (SMC), Transitional MCMC (TMCMC), etc. need to be used to determine the posterior distribution. Among the various simulation methods, random walk MCMC simulation is the most commonly used and is the workhorse of Bayesian inference. However, the disadvantage with using MCMC is that it is a sequential simulation and cannot be completely parallelized. Moreover, as the MCMC simulation is based on a single chain tracking the entire posterior distribution through random walks, the convergence is very slow and requires a large number of simulations of the order of 10^3 to 10^5. As each simulation involves the evaluation of the model, MCMC cannot be applied in case of inference involving large and complex models, whose simulation time is long. The alternative, for large models, as is the case in this thesis, is to use methods which can exploit parallelism to reduce computational overhead, such as Sequential Monte Carlo. Both the methods are implemented in Python and used according to the complexity of the inference problem involved.

4.2.1 MCMC Simulation: Adaptive M-H algorithm

Among the different algorithms to perform random walk MCMC simulation, Metropolis Hastings (M-H) algorithm is one of the most popular and simpler methods [62]. One of the primary performance measure that differentiates the effectiveness of the various MCMC algorithms is the speed at which the algorithm converges to the stationary posterior distribution, i.e the equilibrium distribution which does not change any further with additional samples, especially when the number of parameters involved is large. As the conventional random walk M-H algorithm has a slow rate of convergence, primarily due to the difficulty involved in the choice of an appropriate proposal distribution covariance, adap-

tive Metropolis Hastings algorithm is a preferred choice especially when dealing with a large number of parameters. The adaptive Metropolis Hastings algorithm, similar to the method of Atchade and Rosenthal[63] is used, where the proposal covariance is adaptively tuned based on continuous monitoring of the acceptance rate. The acceptance rate, here, is defined as fraction of accepted samples of the total number of samples that have been proposed during random walk. Instead of tuning the covariance after every simulation, the adaptive tuning is performed at a frequency of w simulations. A general description of the algorithm steps is briefly presented in Fig. 4.1.

MCMC simulation: Adaptive M-H algorithm

1. ***Initialization:*** Initialize chain at sample θ_0 and choose proposal distribution of variance matrix Σ_0 and set $\theta_{curr} = \theta_0$ and $\Sigma_{curr} = \Sigma_0$

2. ***Propagation***: Sample $\theta_{cand} \sim MVN(\theta_{curr}, \Sigma_{curr})$, where $MVN(\theta_{curr}, \Sigma_{curr})$ is the Metropolis Hastings transition kernel.

3. ***Calculate accpetance ratio*** : Calculate posterior distribution $P(\theta_{curr}|y)$ and $P(\theta_{cand}|y)$ based on Eq. 4.1 without evaluating the denominator and compute $\alpha = \min\left(\left(\frac{P(\theta_{curr}|y)}{P(\theta_{cand}|y)}\right), 1\right)$

4. ***Accept/reject***: Generate $u \sim \mathcal{U}(0,1)$ and accept θ_{cand} if $\alpha > u$; reject otherwise.

5. ***Adapt proposal***: Compute acceptance rate, r and repeat the following for every w simulations.

If $r < 0.23$, set $\Sigma_{new} = e_d \Sigma_{curr}$ else

set $\Sigma_{new} = s_d \Sigma_{curr}$ and set $\Sigma_{curr} = \Sigma_{new}$

Figure 4.1: Steps involved in the adaptive M-H algorithm to perform MCMC simulation

In the algorithm, defined in Fig.4.1, the notation MVN refers to the multivariate Gaussian distribution, $\mathcal{U}$ refers to the uniform distribution, and the reference value 0.23 (Atchade and Rosenthall, 2005) for the acceptance rate is the optimal acceptance rate required for a good mixed chain. In the actual implementation of the algorithm, additional tuning is performed depending upon the distance (number of simulations) between the two successive accepted samples. The multiplicative parameters e_d and s_d define the expanding and shrinking rate of the proposal covariance respectively. Depending on the current acceptance

rate, the covariance is either shrunk or expanded by the respective multiplicative parameters. The values of the these two parameters along with w are determined by tuning with trial and error runs until a faster convergence is obtained.

4.2.2 Sequential Monte Carlo (SMC)

Sequential Monte Carlo (SMC), which belongs to the class of particle tracking algorithms, on the other hand, has a more rapid convergence compared to MCMC algorithms. The basic idea of SMC is to start from a simple prior distribution and gradually transition over to the complex posterior distribution via a sequence of annealed intermediate posterior distributions. These sequence of intermediate target distributions are obtained by using either data tempering, wherein the data used in the likelihood are gradually introduced in batches at each iteration or by using likelihood tempering, wherein the likelihood distribution itself is gradually introduced in the posterior as follows.

$$\pi_t(\theta) = \frac{\mathcal{L}(y|\theta)^{\gamma_t} P(\Theta))}{\int_\theta \mathcal{L}(y|\theta)^{\gamma_t} P(\Theta)) d\theta}, \tag{4.2}$$

where γ_t varies from 0 at $t = 0$ to 1 at $t = T$. Here, T is the number of iterations, in terms of tempering steps, required to reach the actual target posterior distribution. In contrast to MCMC, several particles are used simultaneously to track the target distribution. The SMC algorithm is a combination of an importance sampling (a mutation step) and resampling (a correction step). The steps involved in the SMC algorithm are illustrated in Fig.4.2. In mutation, each particle θ_{t-1} from π_{t-1} uses a MCMC chain to move to its target distribution π_t as θ_t . The transition from π_{t-1} to π_t is small in magnitude, and hence short MCMC chains of length between 10-100 simulations are sufficient for the mutation of θ_{t-1} to θ_t. Since the particles have different weights or probabilities in their target distribution compared to their source distribution, a correction step is performed to reassign the weights of each sample according to its likelihood in the distribution. Finally, the

SMC algorithm

1.Initialization: Set $t = 0$, for $i = 1..N_{smc}$, sample $\theta_0^i \sim \pi_0 = P(\theta)$ and set $w_0^i = 1/N_{smc}$

2.Mutation or Prediction: Set $t = t + 1$, sample $\theta_t^i \sim MVN\left(\theta_{t-1}, \widehat{\sigma_{\pi_t}^2} I\right)$, where $MVN\left(\theta_{t-1}, \widehat{\sigma_{\pi_t}^2} I\right)$ is the Metropolis Hastings transition kernel.

$\widehat{\sigma_{\pi_t}^2} = \frac{\sum_{j=1}^{Nsmc} w_{t-1}^j \{\theta_{t-1}^j - \widehat{E_{\pi_t}}\}\{\theta_{t-1}^j - \widehat{E_{\pi_t}}\}}{\sum_{j=1}^{Nsmc} w_{t-1}^j}$; $\widehat{E_{\pi_t}} = \frac{\sum_{j=1}^{Nsmc} w_{t-1}^j \theta_{t-1}^j}{\sum_{j=1}^{Nsmc} w_{t-1}^j}$

3.Correction : The weights of the particles in the target distribution are reassigned via an incremental weight factor as follows: $w_t^j = w_{t-1}^j P(y|\theta_t)^{\gamma_t - \gamma_{t-1}}$ and the normalized weights are given by $\widetilde{w}^j{}_t = \frac{w_t^j}{\sum_{j=1}^{Nsmc} w_t^j}$

4.Selection : Samples are resampled among θ_t according to their normalized weights $\widetilde{w_t}$ and those with negligible weights are discarded. Resampling is done using *multinomial* resampling where samples θ_t are replicated in proportion to their respective weights. Resampling is done only when the Effective Sample Size $(ESS) < ESS_{thresh}$, where $ESS = \sum_j 1/\widetilde{w}_t^{j\,2}$ and $ESS_{thresh} = N_{smc}/2$. The weights of all samples are made equal after resampling as $w_t^j = \frac{1}{N_{smc}}$.

Figure 4.2: Steps involved in the SMC algorithm

samples which lose their weight are discarded by a selection step. A schematic describing the SMC algorithm steps is shown in Fig.4.3. Each of the involved steps of the SMC sampler algorithm along with the choice of MCMC transition kernel is given below. In the steps in Fig. 4.2, the tilde such as in ' $\sim$ ' implies that the samples of random variable are from the probability distribution. Adaptive adjustment of the tempering schedule is done in order to avoid a rapid deterioration of particles as well as to accelerate convergence. The value of γ_t is therefore adaptively chosen according to: $\gamma_t = at^\beta$, where the rate of increase governed by β is reduced when $\mathrm{ESS_t} - \mathrm{ESS_{t-1}} > \mathrm{ESS_{limit}}$. The optimal values of a and β are determined according to some initial studies performed on a smaller model of just the adsorption equilibrium capacity containing the isotherm parameters. The values thus found are a= 0.0005 and β=2.8. If the difference in the ESS between subsequent iterations exceed $\mathrm{ESS_{limit}}$ (= 0.4 $\mathrm{ESS_{thresh}}$), then the value of β is reduced by 0.1 and the parameter is updated as β_{new}=β_{old}-0.1. All the cases reported in this work are run with the number of particles N_{smc} around 100 to 200, the value of which was chosen based on initial Mean

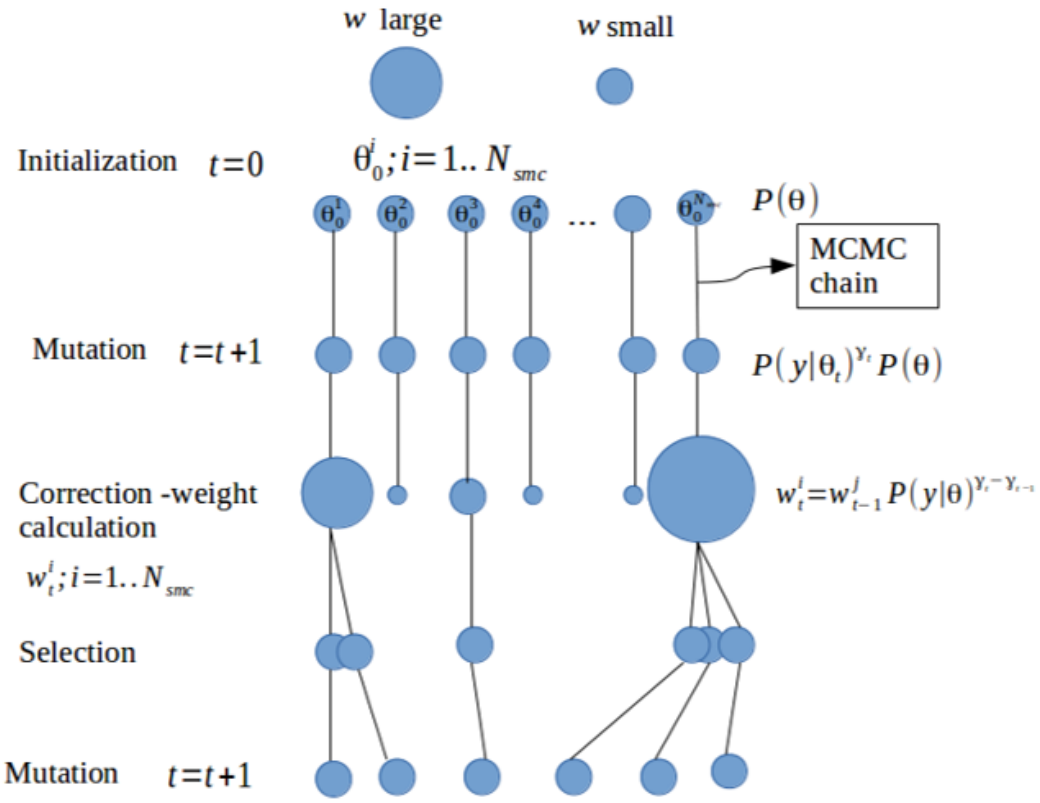


Figure 4.3: Illustration of samples propagation and distribution tracking in SMC

Square Error (MSE) convergence studies performed on the equilibrium adsorption capacity model (Eqs. 2.32 - 2.35).

4.3 Sources of uncertainties and uncertain parameters

Among the model parameters, θ that were estimated using conventional methods of estimation in the previous chapter, uncertainties will be characterized only for the adsorption isotherm, θ_{iso}, and mass transfer parameters, θ_{mass}. The heat transfer parameters h_t and U are assumed to be at their values estimated by conventional methods. The reason to ignore the inference of heat transfer parameters is explained as follows. Due to the small scale of the experimental set-up, almost all of the adsorption heat released is lost to the surrounding heat sinks in the lab-scale set-up, resulting in a very minor temperature rise. Moreover, the convective heat transfer by gas moves the observed thermal front significantly ahead of concentration front, thereby making the adsorption process nearly identical to isothermal adsorption. In other words, the CO_2 breakthrough curve is almost identical to the isothermal adsorption breakthrough curve. Therefore, the sensitivity of breakthrough curve with respect to heat transfer parameters is negligible at the experimental operating conditions. As a result, it is quite difficult to characterize the uncertainties in the heat transfer parameters with the available breakthrough data.

The experimental data that are used to characterize the uncertainties in model predictions are the same CO_2 breakthrough data, C_{br}, as in Figs. 3.5(a)-3.5(b), which were used for experimental model validation in the previous chapter. In addition, the equilibrium capacity measurements, q_{eq}, at various temperatures and pressures (Fig: 3.1), obtained from TGA experiments, are also used together with the breakthrough curves. The sources of uncertainties that are identified in the experimental set-up and thereby in the observed measurements are as follows: a) the measurement noise in the CO_2 breakthrough curve and equilibrium measurements, the standard errors of which are denoted as σ_{C_g} and σ_q for measurements C_{br} and q_{eq} respectively, b) the uncertainty in the placement of hollow

fibers within the module, effectively varying the cross section of gas flow around each fiber, whose outer radius is denoted as r_{fs} in the model, and c) random variations in the loading of amine within the silica support and further in the impregnation of sorbents within the fiber, effectively varying the adsorption isotherm and mass transfer parameters. Besides, there could also be uncertainty around many other input variables in the experimental set-up such as inlet gas volumetric flow rate, inlet gas composition etc. However, these were found to have very minor fluctuations and not to have significant influence on the breakthrough curve. Nevertheless, all the input variable uncertainties that are not explicitly accounted, although minor, are assumed to be absorbed as uncertainties in physical model parameters. Finally, the inadequacy of the model to explain the experimental outcome due to various modeling assumptions, is also included as an additional uncertainty.

4.4 Parametric inference: Various approaches

The goal of uncertainty characterization or parametric inference step, is to find the model with a distribution of parameters, as required, to explain all the experimental breakthrough data (Figs. 3.5(a) - 3.5(b)) and equilibrium data (Fig:3.1) at various operating conditions within the measurement error σ_{C_g} and σ_q respectively. In other words, the model predictions along with the characterized distributions of its model parameters and other uncertain input variables, should fit all the experimental data within its determined error deviation. In order to achieve this goal, four different approaches are used. Three of the approaches assume that the model is adequate and the uncertainties are all with respect to the model parameters or model input variables alone, while the last approach considers that the model itself may be inadequate in addition to parametric uncertainties.

For all the four approaches, the likelihood is evaluated based on considering both the experimental data of $\mathrm{CO_2}$ breakthrough and q_{eq} data simultaneously and is described in the equation below. It is defined as a product of the individual likelihoods, as the $\mathrm{CO_2}$ breakthrough curve is conditionally independent of the equilibrium adsorption capacity q_{eq}

for a given set of isotherm parameters, θ_{iso}.

$$\mathcal{L}(y|\theta) = \mathcal{L}(C_{br}, q_{eq}|\theta) = \mathcal{L}(C_{br}|\theta)\mathcal{L}(q_{eq}|\theta_{iso})$$
$$= \frac{1}{\sqrt{(2\pi)^{n_{cexp}}|\Sigma|}} \frac{1}{\sqrt{2\pi}\sigma_q^{\,n_{qexp}}} \exp\left((\mathrm{C_{br}} - \widehat{\mathrm{C_{br}}})^{'} \Sigma^{-1} (\mathrm{C_{br}} - \widehat{\mathrm{C_{br}}})\right) \exp\left(-\frac{(\mathrm{q_{eq}} - \widehat{\mathrm{q_{eq}}})^2}{2\sigma_{\mathrm{q}}^2}\right) \tag{4.3}$$

By decomposing the overall likelihood, which includes multiple data sets, as a product of individual likelihoods, data from multiple sources and of different scales can be seamlessly integrated into the parametric inference. Such a likelihood decomposition can be achieved by identifying or designing experiments, which can conditionally isolate certain parameters from the rest of the model, as is done for adsorption isotherm parameters. The likelihood in Eq. 4.3 is assumed to be of Gaussian distribution, wherein the errors between experimental data and model predictions are assumed to be additive and are defined as $\epsilon_c = C_{br} - \widehat{C_{br}}$ and $\epsilon_q = q_{eq} - \widehat{q_{eq}}$ for breakthrough and equilibrium capacities data respectively. The variables with hat$\widehat{\ }$denote the model predictions. Herein, ϵ_c denotes the sum of all the residual errors including measurement noise, variability in the experimental conditions, as well as model inadequacy and appear respectively in $\epsilon_c = \epsilon_{noise} + \epsilon_{var} + \delta$. Wherein, the error ϵ_{var} denotes the error due to all other random variability in the data, which are not accounted for in either the model parameters or in the model inadequacy δ. The covariance matrix, Σ in Eq.4.3, is simply a diagonal matrix when the errors are assumed to be uncorrelated having variance $\sigma^2_{C_g}$. On the other hand, in the formulation with model discrepancy discussed below, the matrix has non-zero off-diagonal covariance elements.

With an ideal model that explains the process completely, the error ϵ_c will be at its minimum and equal to the measurement noise ϵ_{noise}. In other words the model, with all the uncertainties characterized within its parameters alone, should be able to predict within the error bars of measurement noise with variance, $\sigma^2_{C_g}$.

4.4.1 Prior distribution

The entire set of parameters including the hyper-parameters, θ_{hyp} to be inferred, is given by the set $\theta = [q_{m0}, \eta, A, B, lnb_0, \Delta H_0, D_{p0}, \alpha, E, \psi, \theta_{hyp}]$. As noted in the introductionary chapter, the parameters that are used to characterize the uncertainty distribution, and are not part of the physical model parameters are referred to as hyper-parameters. In the case of parametric inference without model discrepancy, there are only two hyper-parameters and the set is given by $\theta_{hyp} = [\sigma_q, \sigma_{C_g}]$. Whereas, in the case of model discrepancy, there are four hyper-parameters as given by the set $\theta_{hyp} = [\sigma_q, \sigma_{C_g}, \sigma_{model}, \phi_t]$, the details of which are discussed in the model discrepancy approach section below. There are a total of fourteen parameters at the maximum to be inferred, among which ten of them are the actual model parameters. Inferring all the parameters simultaneously with a vague or less informative prior is quite a difficult problem. To tackle this challenge, the parameters are divided into two different subsets, namely adsorption isotherm parameters, θ_{iso} and mass transfer parameters, $\theta_{mass} = [D_{p0}, \alpha, E, \psi]$, and separate numerical and experimental runs are designed and identified respectively to gather prior information for both the parameter subsets.

The prior distribution for the adsorption isotherm parameters along with their hyper-parameter σ_q is obtained by doing a separate inference problem with the measurements of adsorption capacity, q_{eq} obtained from TGA experiments (Fig.3.1). Bayesian inference was performed as given in Eq.4.4 using MCMC simulation to determine the parametric distribution of isotherm parameters using a mildly informative prior of uniform distribution with large bounds as shown in Table. 4.1. The parameters are scaled down to have similar orders of magnitude, which is helpful when designing the proposal distributions. The parameters are scaled down by factors as are shown in Tables.4.2 and 4.1. The resulting marginal parametric distributions of adsorption isotherm parameters are shown in Fig. 4.4. As the adsorption isotherm model for this inference problem is small and involves only an algebraic system of equations, Bayesian inference via MCMC simulation was computationally

feasible. The parametric distribution, $P(\theta_{iso}|q_{eq})$, obtained from inference is then used as a prior distribution for the adsorption isotherm parameters θ_{iso}, which are later updated with CO_2 breakthrough data during the parametric inference of all parameters θ as given in the equation below (Eq.4.5).

$$P(\theta_{iso}, \sigma_q|q_{eq}) = \frac{\mathcal{L}(q_{eq}|\theta_{iso})P(\theta_{iso})}{\int_\theta \mathcal{L}(q_{eq}|\theta_{iso})P(\theta_{iso})} \tag{4.4}$$

$$P(\theta|C_{br}, q_{eq}) = \frac{\mathcal{L}(C_{br}, q_{eq}|\theta)P(\Theta|q_{eq}))}{\int_\theta \mathcal{L}(C_{br}, q_{eq}|\theta)P(\Theta|q_{eq}))d\theta}. \tag{4.5}$$

With respect to the mass transfer parameters, on the other hand, mildly informative pri-

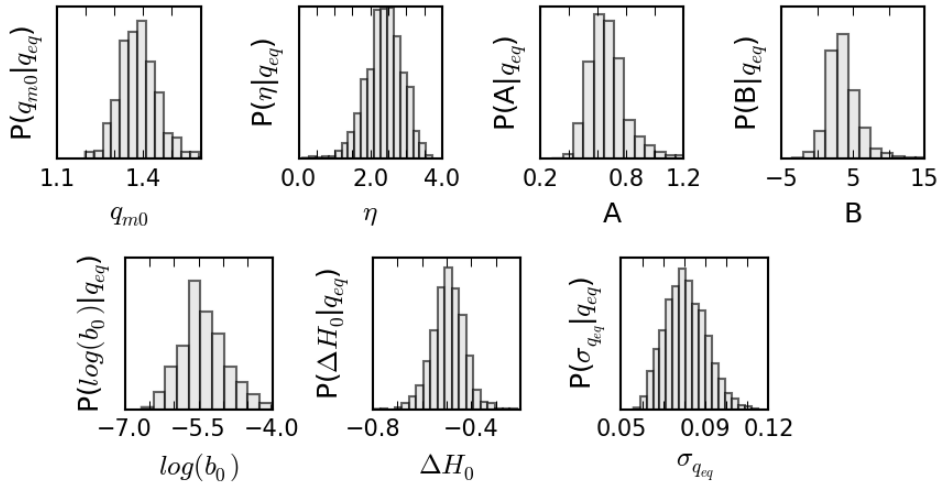


Figure 4.4: Parametric posterior distribution of isotherm parameters based on TGA data, $P(\theta_{iso}|q_{eq})$

ors, such as reasonably bounded uniform distributions are employed. Bounds on the mass transfer parameters were prescreened based on the shape of breakthrough curves obtained using a large number of simulations, performed across a wide range of parameter space of $\theta_{mass} = [D_{p0}, \alpha, E, \psi]$. The extreme range of parameter values, which resulted in a very sharp and unreasonable breakthrough profiles were identified and removed to obtain the

Table 4.1: Bounds of the adsorption isotherm parameters for the uniform prior distribution

Isotherm parameter $\theta_{iso,i}$	Lower bound	Upper bound	Scale
$q_{m0}[mmol/gfib]$	0.1	2.0	$\times$ 1.0
η	0.0	5.0	$\times$ 1.0
A	0.0	3.0	$\times$ 1.0
B	-6.0	30.0	$\times$ 1.0
$log(b_0)[log(\mathrm{bar}^{-1})]$	-9.0	9.0	+log(10^5)
$\Delta H_0[J/mol]$	-2.0	-1.75	$\times 10^5$
Hyperparameter	$P(\sigma_q)$		
σ_q	$\mathcal{N}$(0.09, 4 $\times 10^{-4}$)		

Table 4.2: Bounds of the mass transfer parameters for the uniform prior distribution $P(\theta_{mass})$

Mass transfer parameters $\theta_{mass,i}$	Lower bound	Upper bound	Scale
$D_{p0}[m^2/s]$	0.1	10	$\times$ 10
$E[kJ/mol]$	0.1	20	$\times$ 8314
$\alpha[gfib/mmol]$	0.1	20	$\times$ 1
ψ	0.1	4.0	$\times$ 0.1

Table 4.3: Priors of the hyperparameters P(θ_{hyp})

Hyper-parameters $\theta_{hyp,i}$	P($\theta_{hyp,i}$)
$\sigma^2_{C_g}$	$\mathcal{N}(0.25, 10^{-4})$
σ^2_{model}	$\mathcal{N}(0.25, 10^{-4})$
σ^2_q	KDE of P(σ_q|q_{eq})
σ^2_{var}	$\mathcal{N}(2.1, 10^{-4})$
ϕ_t	$\mathcal{N}(80, 2)$

net bounds on the mass transfer parameters, as shown in Table. 4.2. The prior information for the model discrepancy hyper parameters, obtained using a number of trial simulations performed using different range of values, are also listed in Table.4.3.

4.4.2 Hier-inf: Inference with grouped data series

In the first approach, referred to as 'Hier-inf', separate inference problems are performed with isolated series of data. This is appropriate when it is difficult to obtain a unique set of parameters to explain data under different experimental conditions. Due to random variability associated with sorbent loading in the fibers and other uncertainties in fiber placement within the module, data collected using different modules of even same dimensions and at same experimental conditions can yield different set of parameters, each explaining their respective data alone. In effect, a unique posterior parameter distribution could not be obtained that can fit different series of data within an error of variance $\sigma^2_{C_g}$. This approach is similar to the hierarchical model of Wu et al[64] and 'Hier' model described in Pernot et al [65], which are described as one of the parameter uncertainty inflation (PUI) methods. Whereby, most of the uncertainty is absorbed within the parameters, effectively reducing the model uncertainty. As a result, the prediction uncertainty is largely due to parametric uncertainty alone, besides the random measurement noise. The obtained posterior distributions from individual series inferences are then combined together while performing prediction.

The parametric distribution, $P(\theta_{iso}|q_{eq})$ that was obtained from inference with TGA data (Fig.4.4) is used as a prior distribution while performing the inference of all parameters using breakthrough data as given in Eq.4.5. However, since the parametric distribution, $P(\theta_{iso}|q_{eq})$ obtained from simulation methods (MCMC or SMC), is actually a set of samples occuring in frequency proportional to their probability mass, instead of a pdf, they cannot be directly used as a prior. As, it would be required to evaluate the density of a particular sample chosen while performing the Bayesian inference using MCMC, a proba-

bility density function (pdf) estimate needs to be associated with the distribution (resulting from MCMC or SMC). To do so, kernel density estimation (KDE) method such as Gaussian KDE is used to determine the pdf of $P(\theta_{iso}|q_{eq})$, which is of an arbitrary functional form. KDE is a non-parametric way to estimate the probability density of a function, which does not have a standard functional form such as Gaussian, log-normal, beta distribution etc. The KDE of a distribution, $f(x)$, defined by N set of samples x_i, with $i = 1..N$, is determined as follows:

$$f(x) = \frac{1}{Nh} \sum_{i=1}^{N} \mathcal{K} \left(\frac{x - x_i}{h} \right), \tag{4.6}$$

where $\mathcal{K}$ is the kernel and h is the bandwidth parameter, which is related to the variance of the kernel distribution. For a Gaussian KDE, the kernel is the Gaussian distribution. In other words, Gaussian KDE is obtained by a series combination of several Gaussian distributions, each of which is associated with a sample x_i. The mean of the individual Gaussian distribution is located at the sample value itself x_i, and the variance is defined by the parameter h.

The experimental CO_2 breakthrough data in Figs.3.5(a)-3.5(b), to be used for inference, are collected from two different hollow fiber modules of length 10-inch and 17-inch respectively. In which, the 10-inch module data is collected at varied T_{mod} and 17-inch module is collected at varied Q_{flue}. In this approach, the data that are collected from the same module are grouped together as a series. Accordingly there are two series of data, with which two distinct inferences are carried out (from 10-inch module and 17-inch module) separately.

As the inference involves running the entire hollow fiber model in adsorption step condition, Bayesian inference via MCMC is computationally infeasible and therefore, parallel SMC algorithm is used to perform parametric inference. The SMC simulation for each inference problem is run using 200 particles with complete parallelization by running it across 200 cores in PACE cluster [66] at Georgia Institute of Technology. Both cases take around 28 to 30 SMC pseudo-time iterations, T, to converge by utilizing a total CPU time of around 6 days. Even with complete parallelization of SMC algorithm, through running

each particle simulation separately in a single processor, the computation is still expensive. Given that each iteration takes nearly 1.0 hour to 1.5 hours for completion with the inclusion of a single experimental CO_2 breakthrough curve, and that there are totally four experimental breakthrough curves, each at different conditions, the total time required is nearly 4 hours to 6 hours to obtain the intermediate target distribution $\pi_t = P(y|\theta)^{\gamma_t} P(\theta)$.

The adaptively changing trajectories of the likelihood exponent γ_t, ESS and the value of β, which controls the rate of increase of γ_t are shown in Fig. 4.5 for the case of 10-inch module. As can be seen from the figure, the rate parameter β is decreased by 0.1, whenever there is a large difference between subsequent ESS values, indicating rapid particle depletion. Such a slow increment of γ_t, conditioned on the reduction rate of ESS between iterations, is crucial to obtain a converged posterior distribution. Whenever, ESS falls below the threshold value $\mathrm{ESS_{thresh}} = \mathrm{N_{smc}}/2$, the particles are resampled to facilitate the rejuvenation during the mutation step. The particles are re-sampled again at the end of the final iteration to obtain particles distributed according to the weights distribution, which is easier to handle in post-processing.

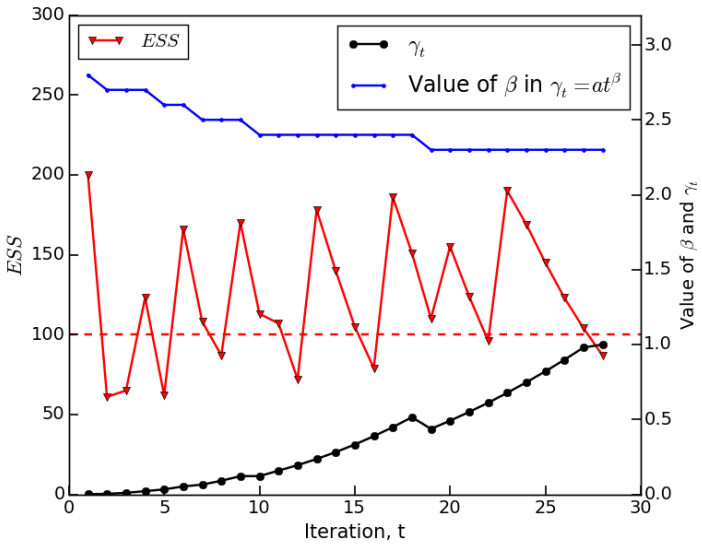


Figure 4.5: Trajectories of Effective Sample Size (ESS), likelihood exponent γ_t and the rate parameter β, which adaptively controls likelihood tempering, for the case with 10 inch module. Dashed line indicates the $\mathrm{ESS_{thresh}}$ value

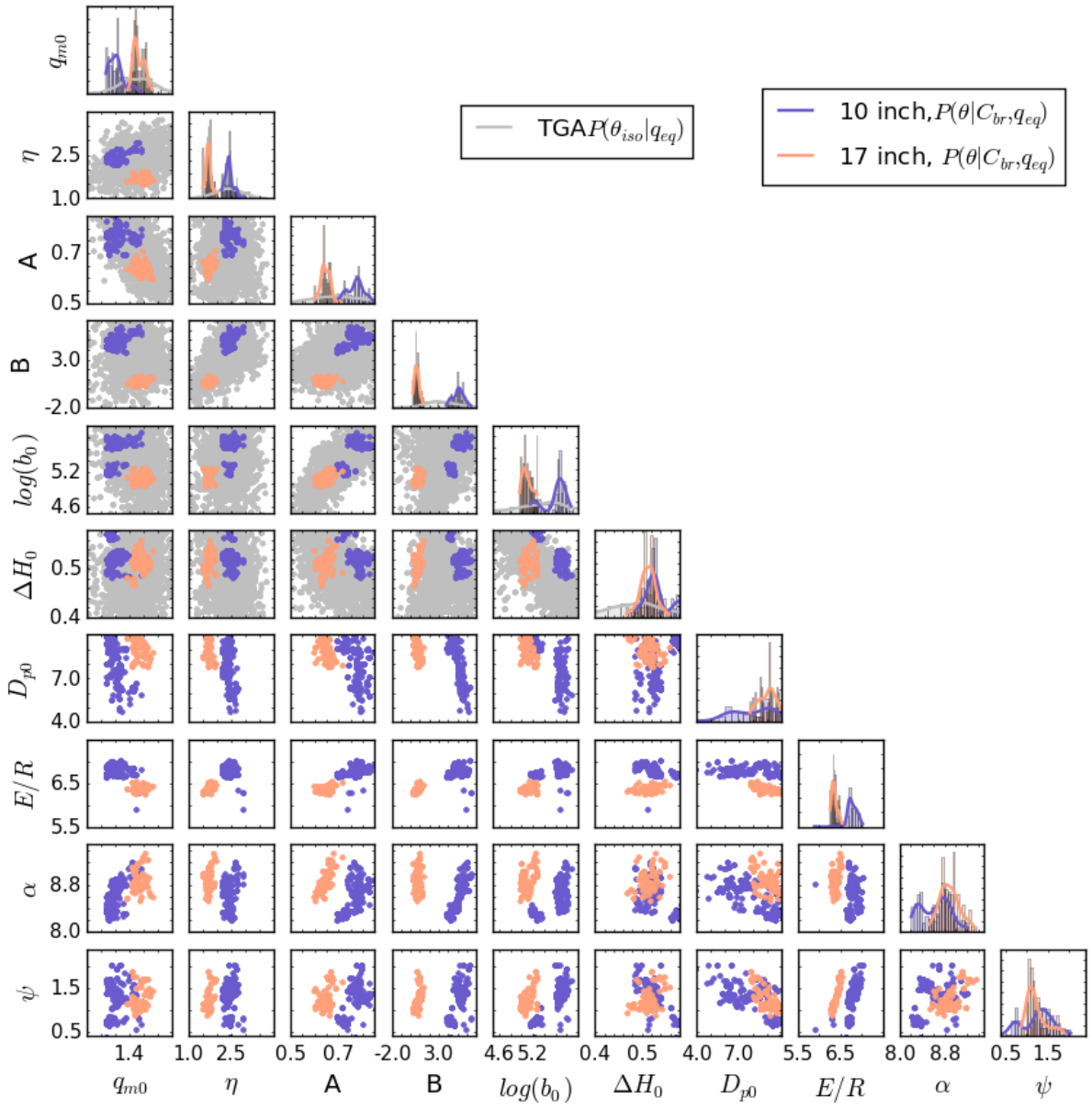


Figure 4.6: "Hier-inf":Posterior parametric distributions $P(\theta_{10}|C_{br-10inch})$ and $P(\theta_{17}|C_{br-17inch})$ obtained with the data collected using two different hollow fiber modules of length 10 inch and 17 inch respectively. $P(\theta_{iso}|q_{eq})$ is the prior distribution of adsorption parameters

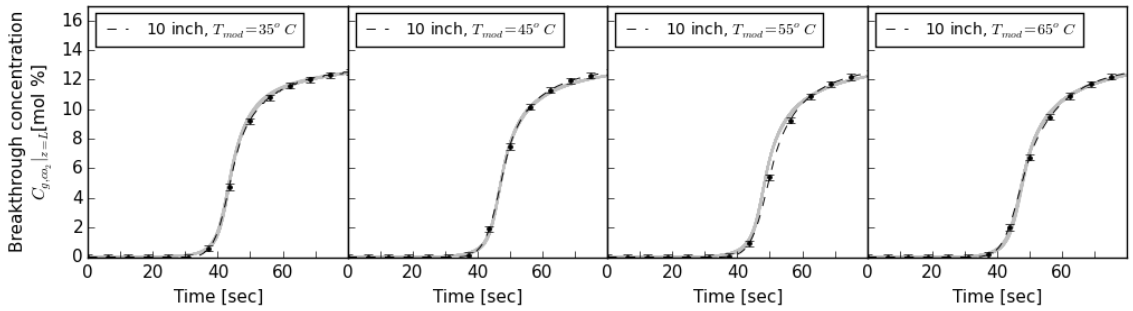
Fig.4.6 shows the posterior distribution obtained individually from both the modules data along with the prior distribution of adsorption isotherm parameters. Although the posterior distribution of both the 10-inch and 17-inch cases yield a very good match with their respective breakthrough data as shown in Fig.4.7, the obtained posterior distributions are very distinct between the two cases. The distinction is observed even with adsorption isotherm parameters, which are not expected to vary significantly with sorbent loading uncertainties that can occur when data are collected using different fiber modules. It can also be observed from Fig.4.6, that the parameters are all highly correlated with one another, which causes the model prediction to have narrow uncertainty band (Fig.4.7), even though the posterior parameter distributions are much wider. The correlation coefficients, $\rho_{X,Y}$, indicating the degree of correlation between each parameter X with respect to the highest correlated parameter Y with respect to each parameter X are listed in Table. 4.4. A value of cross correlation coefficient closer to 1.0 indicates a stronger correlation, and on the other hand, a value closer to 0.0 indicates a weak correlation between parameters X and Y. As seen from the values, ΔH_0 seems to be highly correlated with multiple parameters α and A and similarly ψ has correlation with multiple parameters q_{m0}, B. It could also be the reason for the parameters ΔH_0 and α to have overlapping distributions between the two modules, as it is possible that the least correlated parameters might have converged to different modes while the highly correlated parameter could still retain the same distribution.

4.4.3 Varinflat-inf: Inference with all data-sets and inflated error variance

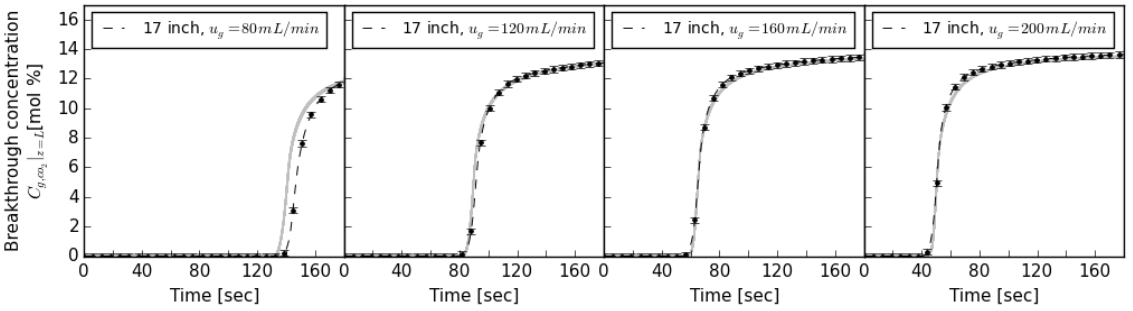
As shown in the previous approach of "Hier-inf", it was found to be difficult to explain all the data with a unique distribution of parameters, while simultaneously assuming that the model is adequate and the only error is of measurement noise with variance $\sigma^2_{C_g}$. Therefore, in order to explain all the data with a unique parameter distribution, the assumptions around uncertainty need to be modified. In this approach, referred to as "Varinflat-inf", the assumption that the only error is of measurement noise is relaxed. Instead, the variance of

Table 4.4: Correlation coefficients for the parameters X with respect to the strongly correlated parameter Y, based on $P(\theta|C_{br,10inch}, q_{eq})$

Parameter, X	$\rho_{X,Y}$ with respect to (wrt) parameter Y
q_{m0}	0.71 wrt ψ
η	0.47 wrt q_{m0}
A	0.70 wrt ΔH_0
B	0.83 wrt ψ
$log(b_0)$	0.61 wrt D_{p0}
ΔH_0	0.85 wrt α
D_{p0}	0.61 wrt $log(b_0)$
E	0.66 wrt α
ψ	0.83 wrt B
α	0.85 wrt ΔH_0



(a)



(b)

Figure 4.7: "Hier-inf":Posterior predictive distributions $P(\widehat{C_{br}}|\theta_s, C_{br-s})$ (s refers to 10-inch or 17-inch accordingly) for (a) 10-inch module at Q_{flue}= 120 mL/min and different T_{mod} and (b) 17-inch module at various Q_{flue} and T_{mod} = 35° C. The posterior mean of standard deviation σ_{C_g} is indicated by the error bars in the curves

the error is increased by adding another variance component σ^2_{var}, which absorbs all the uncertainties due to variability in the data obtained with different modules. All of this variability, which occurs due to variations within the experimental set-up conditions that are not known is termed "residual variability".

The model prediction, accordingly is given by $C_{br-10-inch\,or\,17-inch} = \widehat{C_{br}(\theta)} + \epsilon_{var} + \epsilon_{noise}$, where both ϵ_{var} and ϵ_{noise} are Gaussian uncorrelated errors. The difference, however, is that σ^2_{var} is of higher magnitude than $\sigma^2_{C_g}$ of ϵ_{noise}, so as to absorb all the uncertainties due to "residual variability". The disadvantage, though, is that it is difficult to estimate both the errors separately. Therefore, the total error variance is estimated and the contribution of each error is separated assuming that the measurement noise variance is similar in magnitude to what is estimated in the previous approach of "Hier-inf". This method is also classified as a PUI in Pernot et al[65] and is similar to "Varinf-Rb" approach discussed in their work. The difference, however, is that in the approach of Pernot et al [65], the variance of data is scaled up by a pre-determined factor T. On the other hand, the variance in "Varinflat-inf" approach, is also estimated as needed to fit the data within the total standard error.

The inference problem is run using 200 particles across 200 processors over a period of 14 days. As there are totally eight experimental data of CO_2 breakthrough curves involved at different experimental conditions, each iteration takes nearly 10 hours for completion and the inference was completed in nearly 30 SMC psuedo-time iterations. The posterior parametric distribution obtained from the inference is shown in Fig. 4.8. As the errors are assumed to have a larger variance, it is possible for a larger space of highly correlated parameters to explain and fit the data. As a result, the inferred parametric distributions have a larger variance compared to "Hier-inf", as seen in Fig. 4.8. It can also be noted that most of the parameters have large deviations from a Gaussian distribution. The shift away from Gaussian, inspite of uniform or near Gaussian prior distributions (estimated distributions of isotherm parameters, in Fig 4.4) is due to the high non-linearity of the process model and

the parameters. As seen from Fig.4.9, the unique posterior parametric distribution is able

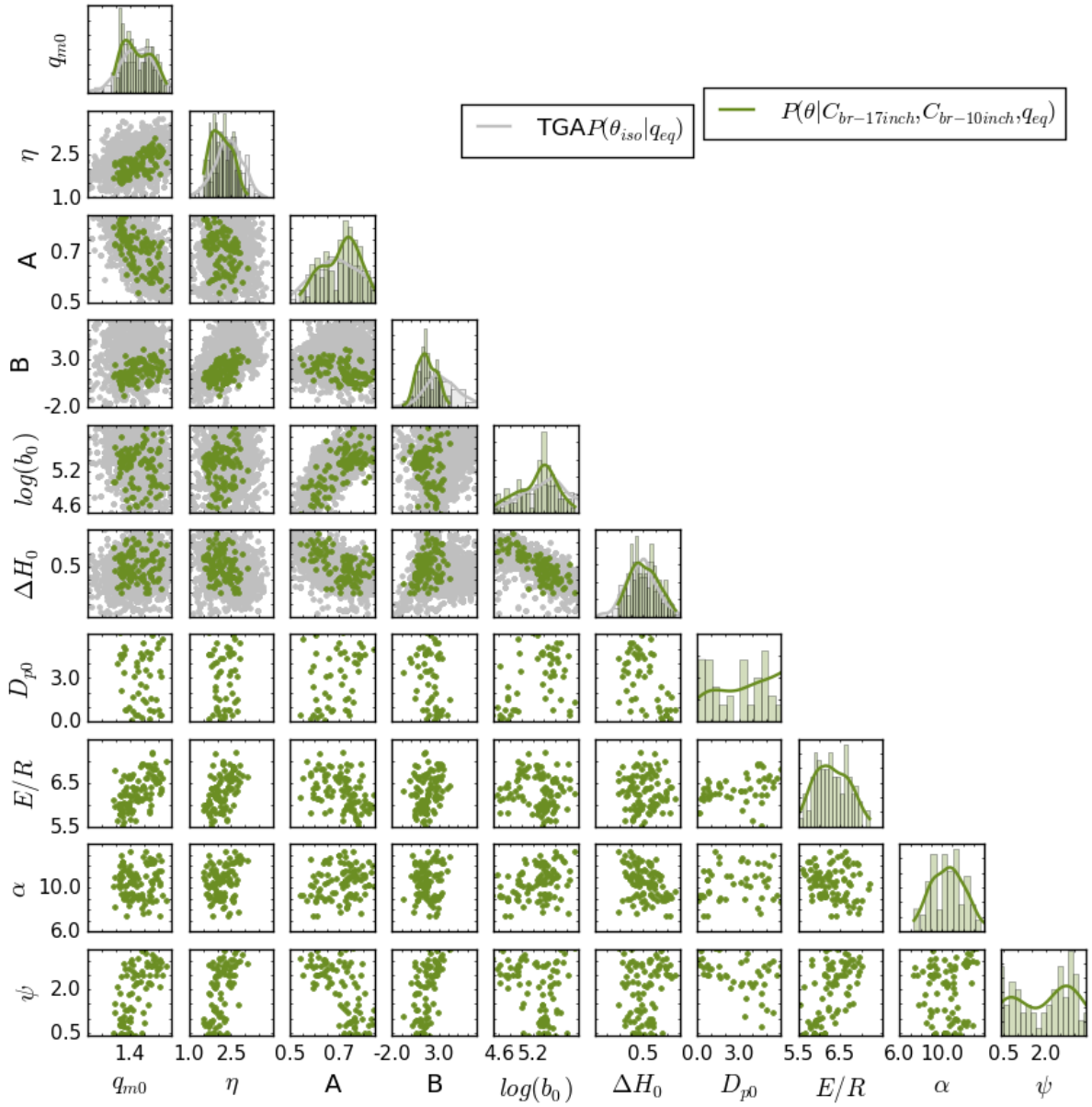
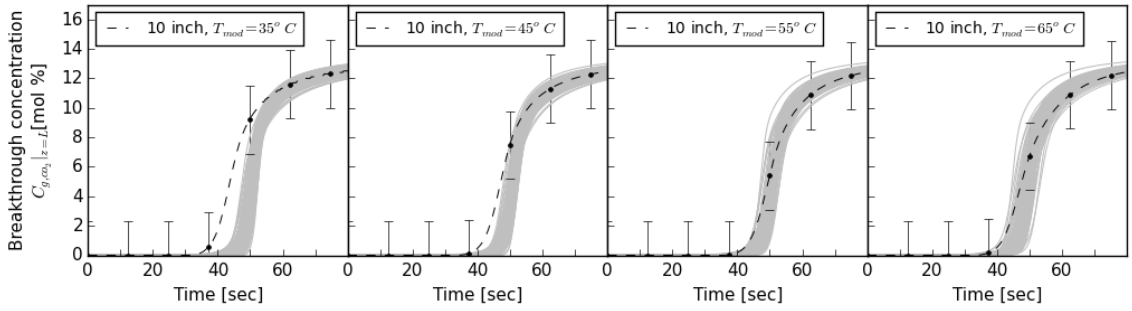


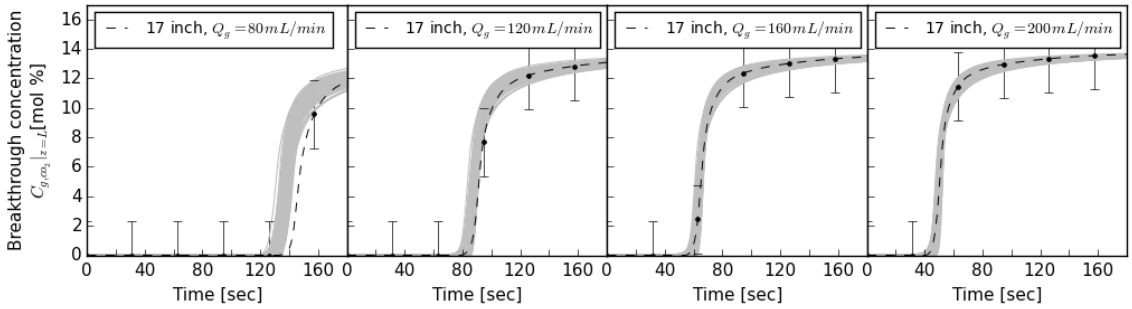
Figure 4.8: "Varinflat-inf":Posterior parametric distribution $P(\theta|C_{br-10inch,17inch})$ obtained using both the modules data simultaneously and considering higher variance σ^2_{var} of error model. Prior distribution from TGA is also shown.

to explain all the data, though, with a larger variance in the model prediction. The other drawback of this approach is that since all the errors, ϵ_{var} are assumed to be uncorrelated and Gaussian, it is possible for the inference to allow for improbable breakthrough curve

shapes, such as a one with a very sharp or a very diffusive shape, which will still be well within the error limit. However, such a prediction did not result in this case, due to the addition of adsorption equilibrium capacity in the likelihood. As the adsorption equilibrium data has lower error variance, it could screen out improbable shapes of breakthrough curves. Although this approach explains all the observed data with a unique posterior distribution, the increased variance in parametric distributions could affect variables other than observed ones to a disproportionate extent. In addition, the prediction uncertainty is also very high in this approach.



(a)



(b)

Figure 4.9: "Varinflat-inf":Posterior predictive distributions $P(\widehat{C_{br}}|\theta_s, C_{br-s})$ (s refers to 10-inch or 17-inch accordingly) for (a) 10-inch module at Q_{flue}= 120 mL/min and different T_{mod} and (b) 17-inch module at various Q_{flue} and T_{mod} = 35° C. The posterior mean of standard deviation σ_{C_g} is indicated by the error bars in the curves

4.4.4 Uresvar-inf: Unfolding the residual variability

Based on the results of the previous two approaches, it is quite evident that it is impossible to obtain a unique parameter distribution that can explain all the data without increasing the error variance, due to the presence of "residual variability" in the data. The alternative is to unfold the residual variability, i.e explicitly account for the uncertainties in the process set-up rather than lumping their effect on the parameters as parametric uncertainties.

In that regard, placement of fibers within the module is expected to be a significant and an influential uncertainty, whose effect is not considered explicitly in the previous two approaches. Each fiber within a module is assumed to be surrounded by an annular hypothetical cross section through which the gas flows. The hypothetical boundary is defined as Happel's free surface radius r_{fs} [50]. Defining a Happel's free surface radius is one of the methods to define the cross sectional area for gas flow around each fiber, which is required to determine the shell side hydrodynamics and mass transfer flux rates. The other method is to use Voronoi tesselation [67], in which each fiber is surrounded by a polygon whose boundaries are defined by perpendicular bisectors of line joining the fiber with its nearest neighbor. Both methods have been used in hollow fiber membrane literature, with each of them having its advantages and disadvantages. One advantage for Happel's free surface approach, though, is its simplicity in modeling as it defines a regular boundary around each fiber. Regardless of the approach used to define the shell side boundary, the randomness of fiber placement within the module can significantly vary the distribution of area and thereby the gas mass and hydro-dynamics between the fibers. As a result, variability in fiber placement has a significant impact on the process output, the CO_2 breakthrough profile.

In this approach, termed as "Uresvar-inf", the uncertainty in the parameter r_{fs} is explicitly accounted for during the inference of parameters θ. With each of the experimental modules of 10-inch and 17-inch lengths having six fibers, the uncertainty of r_{fs} distribution introduces six additional parameters to be inferred resulting in a total of six-

teen physical model parameters apart from two hyper-parameters and is termed as $\theta_{var} = [\theta, r_{fs,1}, r_{fs,2}, r_{fs,3}, r_{fs,4}, r_{fs,5}, r_{fs,6}]$. A single likelihood evaluation, for a single experimental condition, e.g: 120 mL/min in 10-inch module at T_{mod} = 35^{o} C, now involves a maximum of six model simulations, each at a different r_{fs} value. The cross sectional weighted average of the CO_2 concentration at the output that has to be used in the likelihood evaluation of Eq.4.3 for comparison with the experimental data, is computed as follows.

$$\widehat{C_{br}} = \frac{\sum_{i}^{N_{fib}} \widehat{C_{br(r_{fs},i)}} A_{fs,i}}{\sum_{i}^{N_{fib}} A_{fs,i}}, \tag{4.7}$$

where $\widehat{C_{br}(r_{fs,i})}$ is obtained by running the hollow fiber model with $r_{fs,i}$ value of i^{th} fiber. On the other hand, the remaining model parameters which includes θ_{iso}, θ_{mass} and θ_{hyp} are maintained the same between the simulations of all fibers, in which only r_{fs} varies.

Fig.4.10a) shows the numbering of fibers, to be referred further, along with their ideal positions. The prior distribution for all of the six r_{fs} parameters is obtained by giving a random perturbation for each of the fiber from its ideal position. The re-distribution of cross sectional area on perturbation from their ideal positions is done according to the following methodology. Each fiber is grouped with their respective neighbors as shown in Table.4.5. One of the fibers is randomly chosen (termed the chosen fiber), and is perturbed to move closer to any one of its neighbors, which is again randomly selected. For example, as indicated in Fig.4.10 b), fiber no: 6 is moved closer to its neighbor fiber no:1 and in Fig.4.10, fiber no: 3 is moved closer to fiber no: 1. Due to this move, the available symmetric free surface area for both the neighbor fiber and the chosen fiber are reduced. This reduced free space, is therefore, equally re-allotted and added to the other neighbors of the chosen fiber. Accordingly, the free surface cross section of the neighbors of chosen fiber are increased as shown in Figs. 4.10 b) and c). The other possible perturbation that is given is via simulating a symmetric movement of all of the chosen fibers' neighbors simultaneously, by either moving them towards or away from the chosen fiber. For example, the neighbors

of fiber no: 6 are moved towards and away from fiber no: 6 as shown in Figs.4.10 d) and e) respectively. Accordingly, in the former case, where the neighbors move towards the chosen fiber, the reduced free space of chosen fiber is re-distributed equally among all of its neighbors. Similarly, in the latter case, where the neighbors move away from the chosen fiber, the reduced free space of the neighbors are re-distributed and added to the free space of the chosen fiber. Around 10,000 such perturbation simulations are performed and the resulting combinations of r_{fs} configurations of all six fibers are pre-stored and used as the joint prior distribution for all six r_{fs} parameters in the subsequent Bayesian inference. Marginal prior distributions of the r_{fs} values of the six fibers are shown in Fig.4.11.

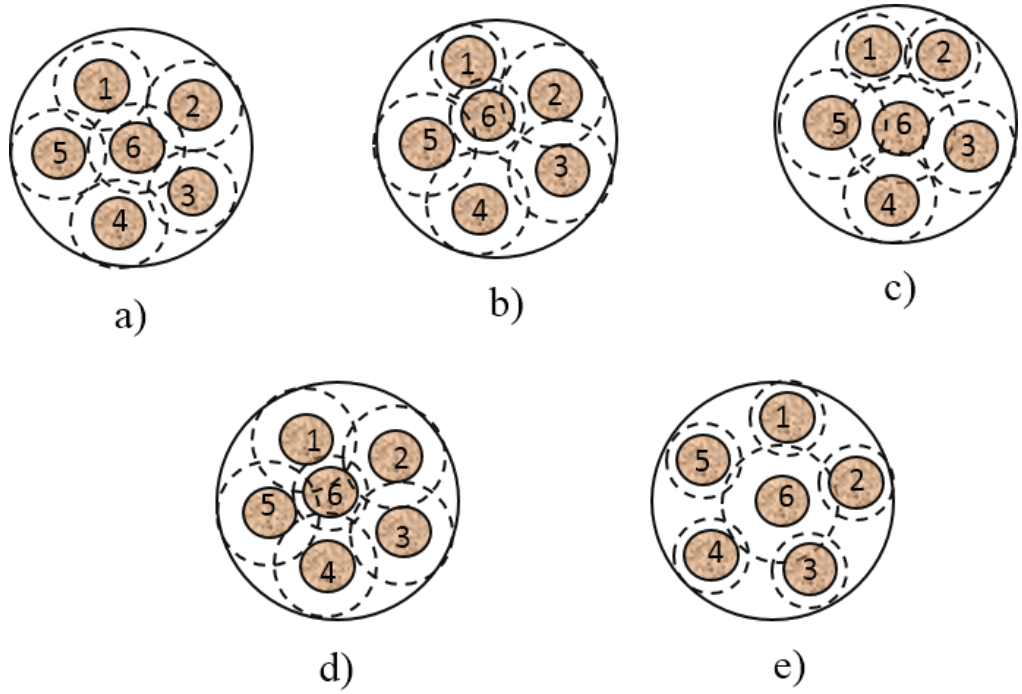


Figure 4.10: "Uresvar-inf": Illustration of r_{fs} variation of fibers within the module when a) all fibers are uniformly placed b) fiber no: 1 is moved closer to fiber no:6 c) fiber no: 2 is closer to fiber no: 3 d) all neigbouring fibers move closer to fiber no:6 e) all neighboring fibers move away from fiber no: 6

As the number of hollow fiber model simulations that are involved for a single evaluation of likelihood is increased by a maximum of six, assuming that there are six distinct r_{fs} parameters, the computational cost is increased by six times. Although these six simula-

Table 4.5: List of neighbors of all fibers in the module, identified according to their numbers shown in Fig.4.10 a)

Fiber no	list of neighbors
1	[2,5,6]
2	[1,6,3]
3	[2,6,4]
4	[5,6,3]
5	[1,6,4]
6	[1,2,3,4,5]

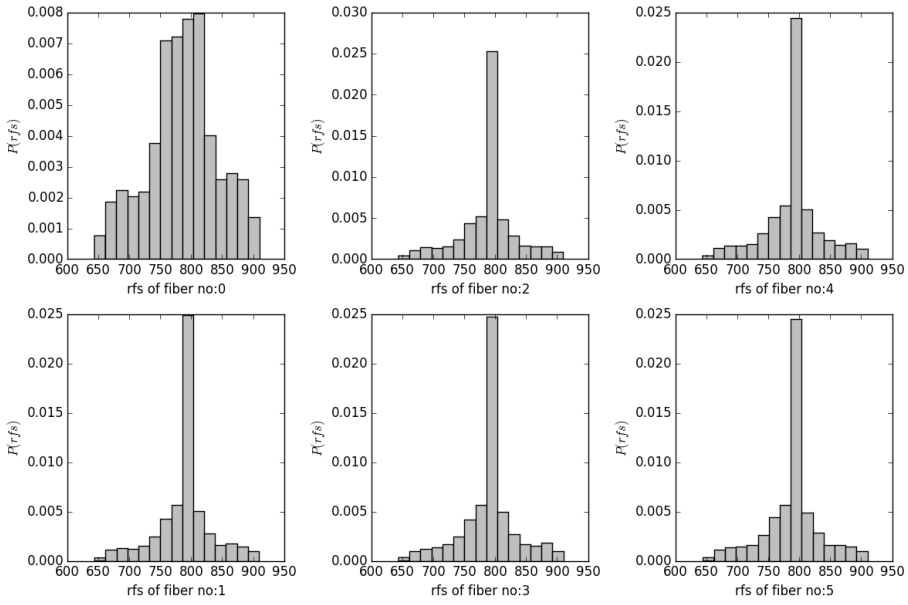


Figure 4.11: "Uresvar-inf": Marginal prior distributions of r_{fs} values of all six fibers

tions can be parallelized, it may require multiple level of parallelization (hyper-threading) to be implemented and non-trivial, given that it needs to be over the existing parallelization, which is across SMC particles. In order to make it computationally efficient without adding another level of parallelization, Sequential Bayesian Updating (SBU) [13] is utilized, whereby the experimental data of each module run is added one by one sequentially in the likelihood, while the prior distribution is updated with each addition and has the information on the experimental data used in the previous run. This is one of the main advantages of the Bayesian approach, which can also speed-up the SMC convergence as the prior information is gradually updated with the data and gets narrower, approaching the final posterior distribution with each update. The SBU principle is described in equations as follows.

$$\begin{aligned} P(\theta|d_1) &\propto \mathcal{L}(d_1|\theta)P(\theta) \\ P(\theta|d_1,d_2) &\propto \mathcal{L}(d_2|\theta)P(\theta|d_1) \\ P(\theta|d_1,d_2,d_3) &\propto \mathcal{L}(d_3|\theta)P(\theta|d_1,d_2), \end{aligned} \tag{4.8}$$

where the available data d_1, d_2, d_3 are added sequentially, with each time using only the latest data (for e.g. d_2) in the likelihood evaluation. Due to the fact that the older data d_1 is encoded in the prior distribution ($P(\theta|d_1)$, which is actually the posterior obtained from the previous run.

Bayesian inference is performed by running SMC algorithm across 120 processors. The eight experimental data are added sequentially using SBU and the posterior parametric distribution obtained is shown in Fig. 4.12. The variance of the r_{fs} values between the fibers in the 10-inch module and in the 17-inch module are shown as the last two parameters in Fig. 4.12. As seen in the figure, the variances of the parameters have reduced drastically compared to posterior parametric distributions of previous two approaches (Fig.4.6, Fig.4.8), partly because only few particles were good enough to explain better all the data that are added sequentially. It could also be because, as there were only fewer particles

across the distribution (120 particles), the variance rapidly decreased with the depletion of particles. In an attempt to recover the diversity in the samples, the variance of the proposal distributions were increased to a large value during the sampling of the intermediate target distributions. This, however, seems to have moved the posterior distributions away from the support of their prior distributions. As can be seen from Fig. 4.12, the posterior distributions, obtained with the 17-inch module for some of the parameters (brown colored) are outside their corresponding priors, which are determined with 10-inch module data (green colored) via SBU. Consequently, the posterior parametric distribution obtained finally does not explain the 10-inch module data runs within the measurement error as seen from Figs. 4.13a). The 17-inch module data, on the other hand, is predicted within the experimental error, indicated by the error bars in the plot (Fig.4.13(b)). In this approach, inspite of adding additional parameters to unfold the residual variability, the posterior parametric distribution could not fit all the data within the measurement error. Therefore, the assumption that the model is adequate is not necessarily correct and therefore, an additional uncertainty in the model structure is introduced.

4.4.5 Mdiscrep-inf: Inference with model discrepancy

As seen from the results of the previous approach, unfolding the "residual variability" could explain only some of the experimental runs, while simultaneously assuming that measurement noise is the only source of error. Therefore, in order to explain for the remaining deviation (Fig.4.7-pack) between the model prediction and the observed experimental data within the measurement error, an additional uncertainty is considered. As, almost all of the anticipated uncertainties are included in the model parameters explicitly, the only other remaining uncertainty that could be added is to the model. Assuming that there exists some phenomenon, which is why the model is unable to explain all the data within the low measurement error, a model discrepancy term is added to the model prediction.

Model discrepancy term δ, in general, accounts for all the systematic errors that oc-

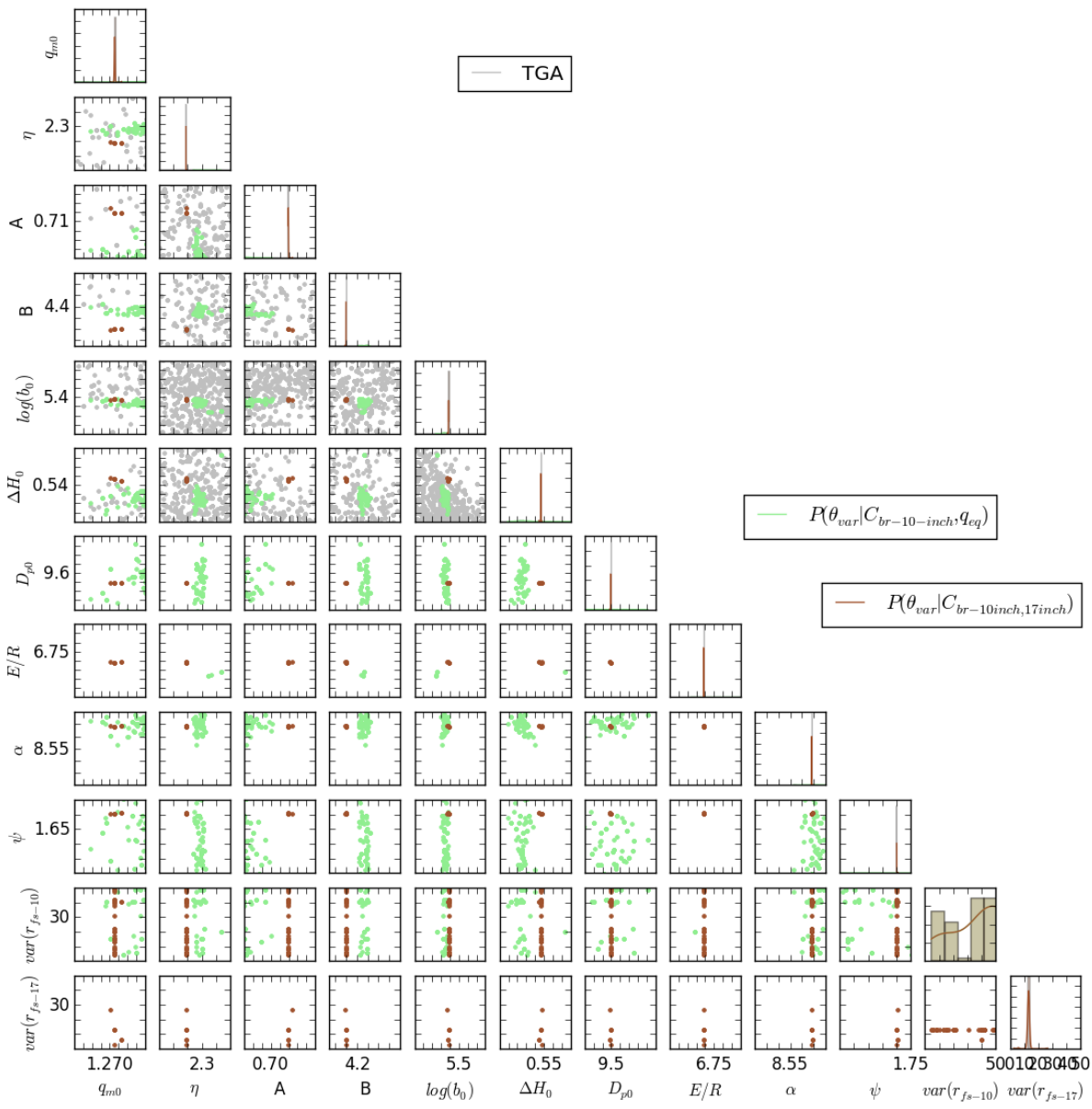
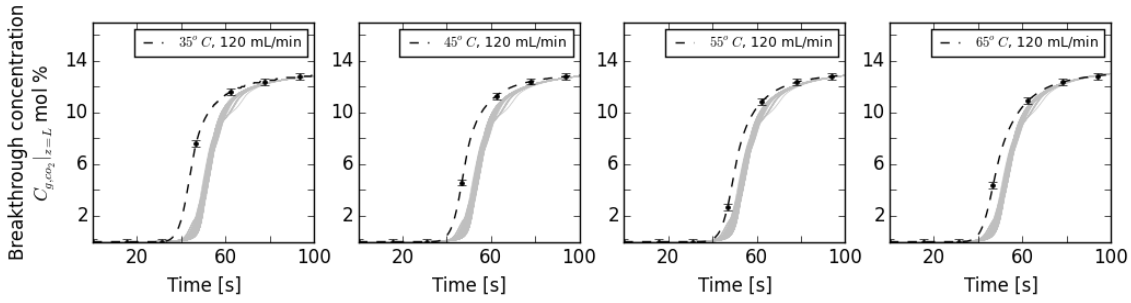
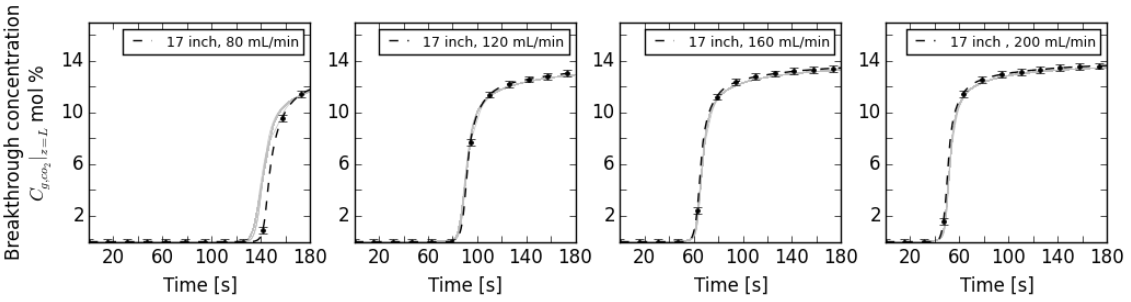


Figure 4.12: "Uresvar-inf":Posterior parametric distributions $P(\theta_{var}|C_{br-10inch,17inch})$ obtained with the data collected using two different hollow fiber modules of length 10 inch and 17 inch respectively. $P(\theta_{iso}|q_{eq})$ is the prior distribution of adsorption parameters



(a)



(b)

Figure 4.13: "Uresvar-inf":Posterior predictive distributions $P(\widehat{C_{br}}|\theta_{var,s}, C_{br-s})$ (s refers to 10-inch or 17-inch accordingly) including packing irregularity uncertainty for (a) 10-inch module at Q_{flue}= 120 mL/min and different T_{mod} and (b) 17-inch module at various Q_{flue} and T_{mod} = 35° C. The posterior mean of standard deviation σ_{C_g} is indicated by the error bars in the curves

cur between model prediction and experimental observation, which could be either due to some physical phenomenon incorrectly described in the model or due to some variability in the experimental setup [10]. It is also important to explicitly account for the model discrepancy, especially in cases where the response variable is continuous, such as in the case of CO_2 breakthrough, which is continuous in temporal dimension. In such cases, an assumption of uncorrelated errors with a large variance, σ^2_{var}, to account for model inadequacies will result in huge random aphysical fluctuations along the continuous variable (time dimension) space, while quantifying prediction uncertainty. The most commonly used approach to capture the model discrepancy is Kennedy O' Hagans methodology [10]. It describes the model discrepancy as a Gaussian process, a multivariate normal distribution as $\delta = MVN(0, \Sigma^m)$, in which the model covariance matrix, Σ^m is defined for the given data as follows.

$$\Sigma^m_{ik} = \sigma^2_m \exp\left(-\frac{(\mathrm{t_i} - \mathrm{t_k})^2}{{\phi_\mathrm{t}}^2}\right) \tag{4.9}$$

Here $[\sigma_m, \phi_t]$ is the additional set of hyper-parameters where σ^2_m is the variance parameter, controlling the spread of systematic bias from its mean and ϕ_t is the range parameters controlling the extent of correlation between two points based on the magnitude of their distance from one another for the variable time, t. The assumption here, is that the all the experimental data has the same measure of model discrepancy in their prediction, i.e same magnitude of model discrepancy is added to the model prediction, regardless of the operating conditions such as different T_{mod} or Q_{flue}. However, in reality, the model inadequacy may be more pronounced for some operating conditions compared to others. Such a modeling with varying model discrepancy along the operating conditional space, require additional hyper-parameters to be added and inferred, which increases the dimensionality of the problem further.

Based on the prior knowledge of the system behavior, the model discrepancy is expected to be negligible before the breakthrough as well as towards the end of adsorption. Such a dynamic variation in the model discrepancy term is achieved by defining the variance

parameter σ_m as a function of the breakthrough curve slopes as follows.

$$\sigma_m = \sigma_{model} \frac{1}{2} \left(\frac{\partial C_{g,co_2}}{\partial t}|_{t=t_i} + \frac{\partial C_{g,co_2}}{\partial t}|_{t=t_k} \right), \tag{4.10}$$

where σ_{model} is the constant variance parameter for the model discrepancy. The total covariance matrix σ, which also includes the measurement noise variance, is therefore defined as $\Sigma = \Sigma^m + \sigma^2_{c_g}$. Recent work by Mebane et al. [24] has used the same model discrepancy formulation of Kennedy O' Hagan [10] to capture the bias in the modeling of thermodynamic CO_2 uptake curves. Their formulation was, however, on a simpler system involving fewer parameters.

Obtaining a good prior distribution for the involved hyper-parameters is a non-trivial task and requires multiple trial simulations to identify the range of values for which the model discrepancy term across the time incorporates the necessary correlation effects. The inference is performed for all the parameters in total using the SMC simulation with 100 particles running across 100 processors. The inference takes nearly two weeks for completion and the posterior distribution of the parameters obtained is shown in Fig.4.14. The predictions of breakthrough curves the experimental data used, including the model discrepancy, are shown in Fig. 4.15. Although the error bars are very small in magnitude, the uncertainty in prediction is much broader due to the presence of model uncertainty. The actual model prediction, excluding the model discrepancy will be identical to what is observed in Fig.4.13. The addition of model discrepancy, thus can explain and fit all the data, however with an additional error term (the model discrepancy) over the measurement noise error.

Based on the results from the four different approaches, it seems highly unlikely to obtain a unique posterior parametric distribution within a low residual error (measurement noise). Either, the variance in the parameters needed to be increased or an additional uncertainty in the model has to be included to fit all the data. The uncertainty estimate of

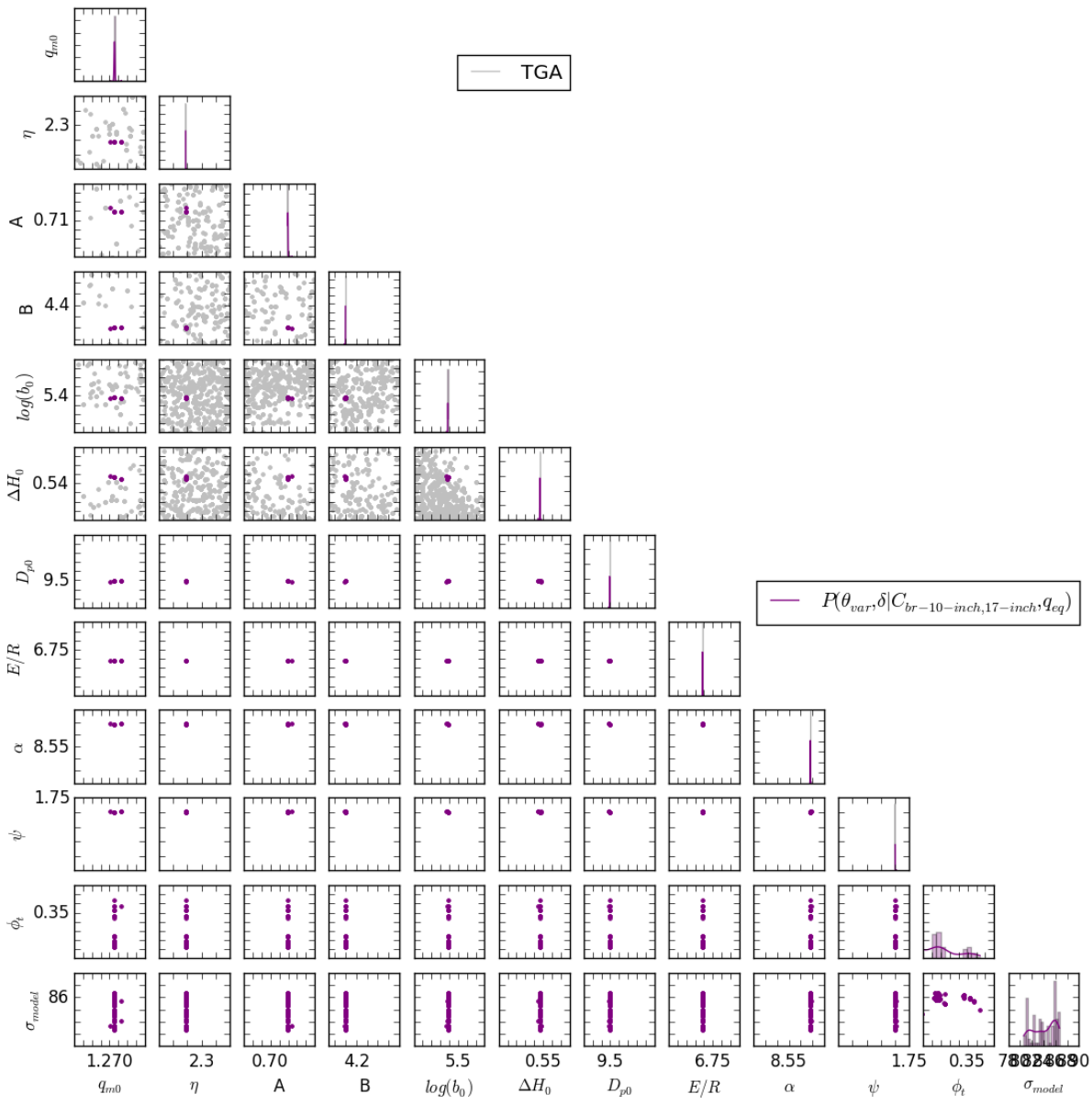
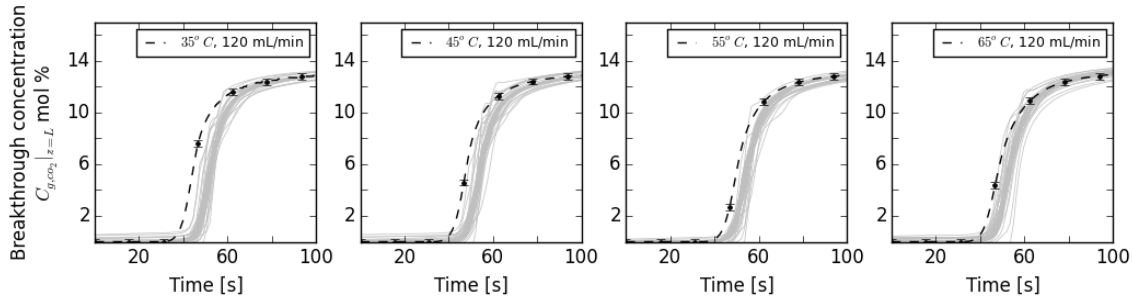
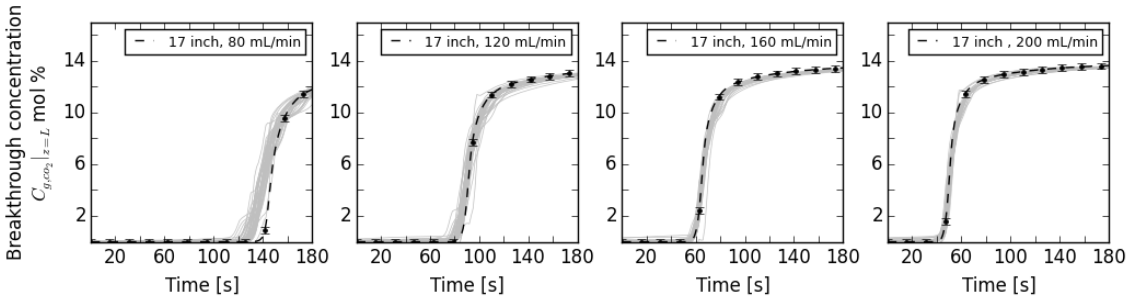


Figure 4.14: "Mdiscrep-inf":Posterior parametric distributions $P(\theta|C_{br-10inch,35^oC})$ obtained with the data collected using hollow fiber modules of length 10 inch, including model discrepancy. $P(\theta_{iso}|q_{eq})$ is the prior distribution of adsorption parameters



(a)



(b)

Figure 4.15: "Mdiscrep-inf":Posterior predictive distributions $P(\widehat{C_{br}}|\theta_{var,s}, C_{br-s})$ (s refers to 10-inch or 17-inch accordingly) including packing irregularity uncertainty for (a) 10-inch module at Q_{flue}= 120 mL/min and different T_{mod} and (b) 17-inch module at various Q_{flue} and T_{mod} = 35° C. The posterior mean of standard deviation σ_{C_g} is indicated by the error bars in the curves

Table 4.6: Comparison of parametric uncertainties with their 99% credible interval [LB,UB] among the four approaches to characterize uncertainties along with the 95% confidence interval [LB,UB] of parameter estimates from conventional methods.

Method	q_{m0}	η	A	B	$logb_0$	ΔH_0
"Hier-inf"	[1.27, 1.43]	[1.49, 2.94]	[0.55, 0.81]	[0.24, 6.42]	[-4.97, -5.87]	[-0.47, 0.59]
"Varinflat-inf"	[1.28, 1.47]	[1.37, 2.86]	[0.56, 0.81]	[-0.49, 4.85]	[-4.63, -5.93]	[-0.41, -0.59]
"Uresvar-inf"	[1.27, 1.28]	[2.18, 2.19]	[0.719, 0.72]	[3.98, 4.00]	[-5.37, -5.38]	[-0.54, -0.55]
"Mdiscrep-inf"	[1.27, 1.28]	[2.18, 2.19]	[0.719, 0.72]	[3.98, 4.00]	[-5.37, -5.38]	[-0.54, -0.55]
"Conventional"	[1.23, 1.43]	[1.27, 2.59]	[0.387, 0.625]	[1.55, 5.31]	[-1.89,0.73]	[-0.54, -0.77]

Approach	D_{p0}	E	α	ψ
"Hier-inf"	[7.80,10.0]	[5.90, 7.02]	[8.20, 9.35]	[0.37, 2.04]
"Varinflat-inf"	[0.30, 9.36]	[5.69, 7.24]	[7.18, 14.15]	[0.42, 3.91]
"Uresvar-inf"	[9.48, 9.49]	[6.69, 6.691]	[8.64, 8.643]	[1.70, 1.704]
"Mdiscrep-inf"	[9.48, 9.49]	[6.69, 6.691]	[8.64, 8.643]	[1.70, 1.704]
"Conventional"	[0.53, 20.25]	[4.80, 7.50]	[6.94, 7.28]	[N/A]

parameters, obtained from the four different approaches that are used to characterize the uncertainties, are summarized via their 99% credible intervals in Table.4.6. The confidence intervals of parameters, obtained from conventional methods, discussed in chapter 3 (Table.3.1 and 3.3) are also included in this table for comparison. Although the credible interval obtained from uncertainty characterization and the confidence interval from conventional methods seems to overlap for some parameters (q_{m0},η and ΔH_0), the credible intervals are wider, and sometimes away from the corresponding confidence intervals from the conventional methods, for most parameters. It could be due to the normality assumption of parameters in conventional methods, and also due to the fact that the estimated parameters were not good enough to fit all the experimental data (Figs.3.5(b), 3.5(a)).

A comparison of key features among the four different approaches, including the total number of parameters involved to be inferred, number of additional hyper-parameters and the computational effort is provided in Table.4.7. As can be seen, both "Hier-inf" and "Varinflat-inf" approaches involve minimum number of parameters and comparable computational cost compared to the "Uresvar-inf" or the "Mdiscrep-inf" approaches.

This chapter characterizes the various uncertainties that occur in the RTSA process model by adopting four different approaches. Once the uncertainties are characterized, the

Table 4.7: Comparison of various features among the four approaches to characterize uncertainties

Approach	Total no of model parameters	No of hyper parameters	Computational cost [hrs]
”Hier-inf”	10	3	37 $\times$1.5$\times$8.0 = 444
”Varinflat-inf”	10	3	30 $\times$1.5$\times$8.0 =360
”Uresvar-inf”	10+12 (12 r_{fs})	3	25 $\times$5.0$\times$8.0 = 1000
”Mdiscrep-inf”	10+12	5	25 $\times$5.0$\times$8.0 = 1000

next step in UQ analysis is to propagate the characterized parametric distribution through the cyclic RTSA model to determine the impact of the uncertainties on the key process outputs such as CO_2 purity and recovery, which is the focus in the following chapter.

CHAPTER 5

UNCERTAINTY PROPAGATION AND UQ SUMMARY

One of the main goal of uncertainty analysis is the uncertainty propagation, in which the uncertainties that are characterized with respect to the model parameters and the model structure, are propagated through the model and their effect on the model observables are quantified. As a result of propagating the rigorously obtained parametric probability distribution conditional on the observed experimental evidence, the uncertainty in model predictions are also rigorously quantified. In other words, uncertainty associated with the future observation of a process variable can be rigorously determined.

As the parameters and the model uncertainties are characterized as a probability distribution via Bayesian inference, the propagation of these uncertainties through the model also yields a probability distribution of the prediction variables. Again, Bayesian inference can be used to determine the probability distribution of prediction variables as follows:

$$P(\hat{y}|y) = \int_{\theta} P(\hat{y}|\theta)P(\theta|y)d\theta. \tag{5.1}$$

In the above equation $\hat{y}$ is just the model prediction alone of an observable without including the calibrated observation error. The term, $P(\hat{y}|\theta)$ is the prior predictive distribution which is characterized by the uncertainty in the parameters θ. The right hand side term, the posterior predictive distribution $P(\hat{y}|y)$ is obtained by integrating the prior predictive distribution over the posterior parametric distribution $P(\theta|y)$. The prediction of the future observation, $\tilde{y}$ can be determined as $\tilde{y} = \hat{y} + \epsilon(\sim \mathcal{N}(0, \sigma_y^2))$, wherein the characterized

observation error ϵ is also included. The observation errors during prediction of future observation are typically assumed to have the same distribution as those of model mismatch (residual) errors that are calibrated during the inference of parameters with the existing experimental data [60]. Accordingly, the future observation errors are also assumed to be uncorrelated Gaussian errors of variance σ_y^2, which was determined during parametric inference with the rest of model parameters.

5.1 Monte Carlo propagation

The posterior predictive distribution of process variables of interest are determined by Monte Carlo propagation of posterior parametric distribution through the model. The posterior parametric distribution, that is obtained from Bayesian inference using simulation methods (MCMC or SMC) is actually a set of samples occurring in frequency proportional to their probability mass. Since, the evaluations of the process model, corresponding to each sample value, θ in the posterior parametric distribution $P(\theta|C_{br}, q_{eq})$, are independent of one another, the propagation can be completely parallelized. As a result of propagation, the model predictions are also obtained as a set of samples occurring in frequency proportion to their probability mass. With the obtained distributions of model prediction variables, the Quantification of Margins of Uncertainty (QMU) can be summarized by determining the credible intervals. The credible intervals with certain α $\%$, in contrary to the classical statistical confidence interval, indicates the interval within which the parameter or the prediction lies with that specific value of probability (=α). Therefore, one can clearly identify the worst and best case scenarios with certain probability (as defined in introductionary chapter) that can be expected from the process, in the presence of the various uncertainties, using the credible intervals of model predictions.

5.2 Uncertainty in cycle performance metrics

The propagation is performed for a larger module with fibers of dimensions L_{fib} = 3 m, r_i = 160 μ m and r_o = 600 μ m. It is to be noted that the module used to collect experimental data during the characterization step (chapter 4), however, is of smaller scale with six fibers of dimensions L_{fib} = 10-inch and 17-inch, r_i = 275 μ m and r_o = 625 μ m. The key assumption during propagation for the prediction of process performance in a larger scale module is that all fibers within the large scale module are identical. The packing irregularity that is considered in "Uresvar-inf", however, is with respect to incorporating the effect of uncertainty in fiber packing for the smaller scale experimental module in order to characterize the parametric uncertainties. However, during propagation, this uncertainty in fiber module packing is ignored. It is to be noted that the uncertainty in the parameter, r_{fs}, representing the packing uncertainty, will not be of the same distribution as was characterized for a smaller scale module. It could rather be characterized by assuming that it follows any standard truncated distributions, such as truncated Gaussian or truncated log-normal distribution etc, since infinite support (or tails) containing infeasible values of r_{fs} need to be truncated. However, considering that additional uncertainty in the input parameter, r_{fs} during propagation means simulating the enire module, which typically includes 100,000 of fibers. This implies that, to determine the prediction uncertainty including the packing irregularity in a larger module, the hollow fiber cyclic model needs to be simulated for all fibers within the module, and for each set of model parameter values θ_i contained in posterior parametric distribution. With that, a single outcome from the module based on a parameter set θ_i, involves averaging the concentration profiles across all the fibers as follows:

$$\widehat{y_{co_2}(\theta_i)} = \int_{A_{fs}} \widehat{y_{co_2}(\theta_i, r_{fs})} P(r_{fs}) dA_{fs}, \tag{5.2}$$

wherein $\widehat{y_{co_2}(\theta_i)}$ is the module prediction of the CO_2 molar composition at the outlet during the cycle corresponding to a single parameter value θ_i and the integration is performed over

the entire fiber module cross section. Here $A_{fs} = 3.14(r_{fs}^2 - r_o^2)$, the Happel's free space cross section around each fiber. In order to determine the complete posterior predictive distribution $P(\widehat{y_{co_2}}|C_{br}, q_{eq})$, however, the computation in Eq.5.2 needs to be performed for every sample θ_i in posterior parametric distribution $P(\theta|C_{br}, q_{eq})$. Although, the distribution $P(r_{fs})$ can be approximated with fewer number of samples (around $\sim$ 50 compared to 100,000), the computation in Eq.5.2 still is required to be done for 50 individual fibers, each with distinct value of r_{fs}. As the computation requires hyper-threading of parallel computation (parallel within parallel processing), it is beyond the scope of the thesis and can be pursued in the future.

The key process variables that are of interest in determining the uncertainty in predictions are the molar fraction of CO_2 at the fiber exit, CO_2 swing capacity, CO_2 recovery and purity. The distribution in water temperature at the exit is also determined. The cycle step times are retained at the same values as used for the simulation with point estimates of parameter values (as in chapter 3). Thereby, the obtained uncertainty in model predictions can be compared to those values obtained via conventional methods. The simulations are run until the cylic steady state (CSS)is reached and the model predictions are reported at the CSS in the posterior predictive distributions.

5.2.1 Hier-inf - UQ summary

As the "Hier-inf" approach resulted in two distinct parametric posterior distributions (Fig.4.6) through inference of individual module data series, both the posterior distributions are propagated through the model and the predictive distributions are obtained. Both the parametric posterior distributions are given equal probability while determining the predictive posterior distributions. Alternatively, one could choose to believe in one set of parametric distribution compared to other, while determining the predictive distribution. The predictive distributions of CO_2 molar composition at the exit are plotted in Fig. 5.1, which has two distinct distributions overlaying one over the other. The predictive distribution of water

temperature at the exit is shown in Fig.5.2. The process performance metrics summary,

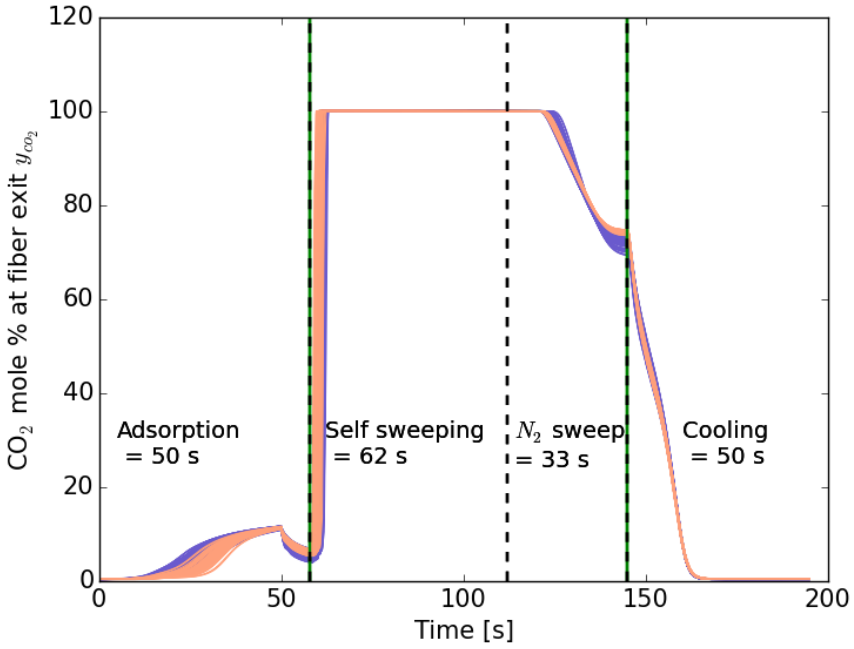


Figure 5.1: "Hier-inf": Model prediction uncertainty in the exit CO_2 molar percent during the cycle at cyclic steady state (CSS)

viz: the predictive distributions of CO_2 swing capacity, product CO_2 recovery and purity are shown in Figs. 5.3, 5.4(a) and 5.4(b) respectively. In contrast with the values of performance metrics reported using point estimates of parameters in chapter 2, the model predictions obtained by incorporating all the uncertainties in experimental data, does indeed have a considerable uncertainty in the values of performance metrics. For example, CO_2 purity can be expected to be anywhere between [91.5%, 96.7%] with 99% probability, based on all the experimental evidence. On the other hand, the purity is predicted to be 95% alone via the conventional methods. All the uncertainty intervals reported in this chapter are of 99% confidence levels, unless otherwise explicitly mentioned.

It is to be noted that these predictive distributions refer to the model predictions alone, without considering the observation error that is expected in reality. The actual future predictive distributions will also include the uncorrelated Gaussian error, ϵ_{noise}, characterized as $\mathcal{N}(0, \sigma^2_{C_g})$, where σ_{C_g} = 0.25. For example, the future prediction of outlet CO_2 mole

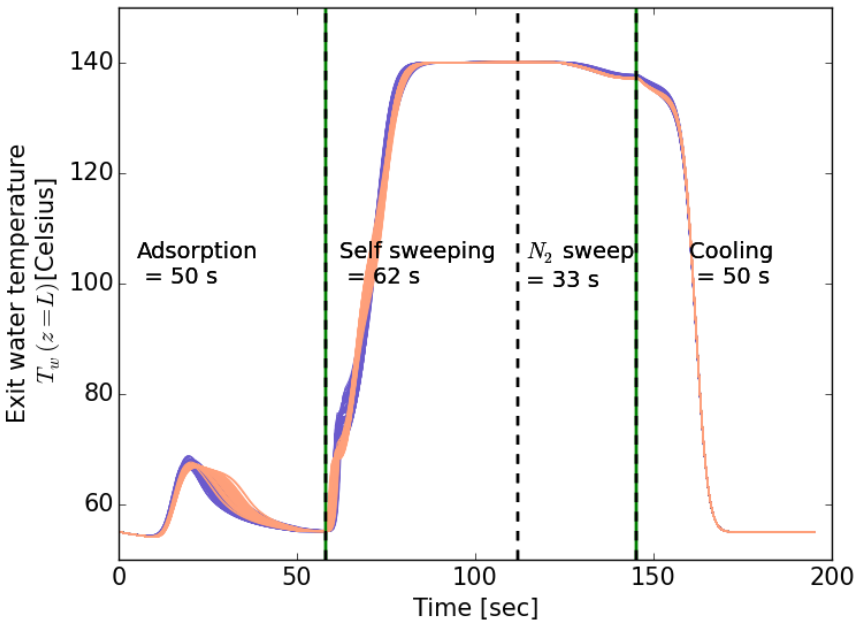


Figure 5.2: "Hier-inf": Model prediction uncertainty in the exit water temperature T_w during the cycle at CSS

fraction will be $y_{\tilde{CO_2}} = \widehat{y_{CO_2}} + \epsilon_{noise}$. As the standard error, σ_{C_g} in this approach is negligible, the future predictive distribution will be only marginally different from the model predictions reported in Figs. 5.1 - 5.4(b).

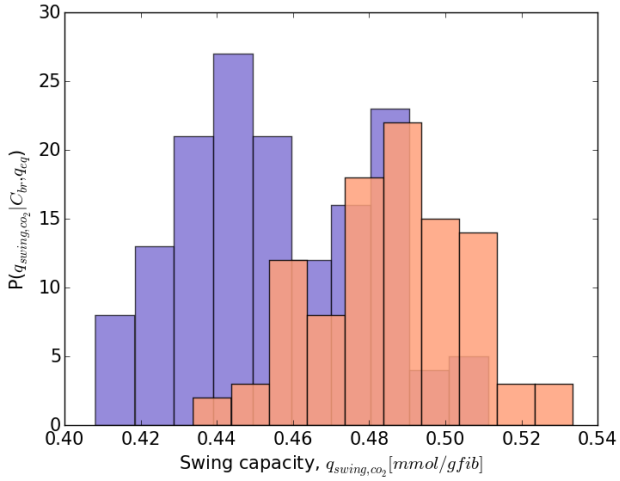
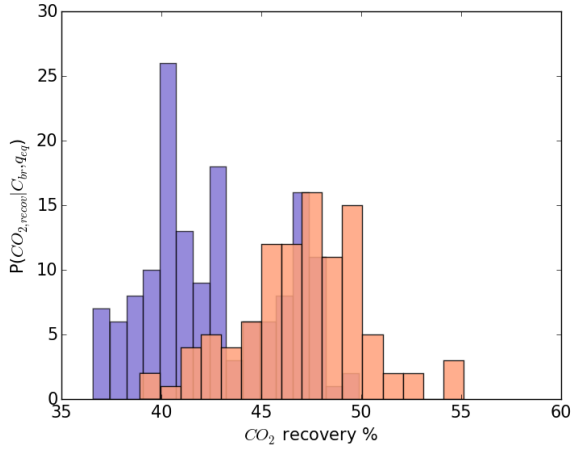
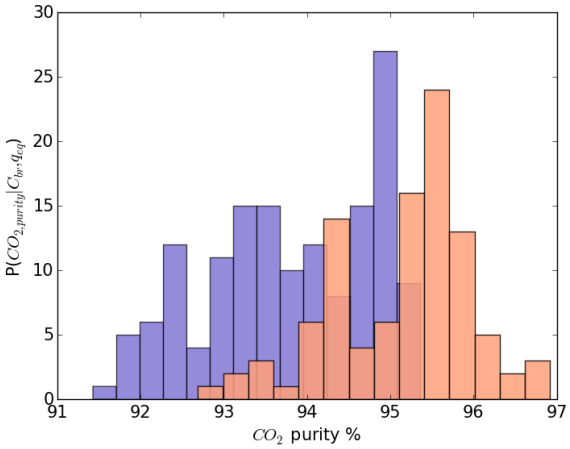


Figure 5.3: "Hier-inf": Model prediction uncertainty in the CO_2 swing capacity at CSS



(a)



(b)

Figure 5.4: "Hier-inf": Model prediction uncertainty in a) the CO_2 recovery and b) CO_2 purity at CSS

5.2.2 Varinflat-inf - UQ summary

The posterior parametric distribution from "Varinflat-inf" approach, shown in Fig. 4.8, is also propagated through the model and the posterior predictions of various process variables of interest are obtained. Figs.5.5 and 5.6 show the CO_2 molar fraction at the outlet during the cycle at CSS. As the parameters have much wider posterior distributions in this

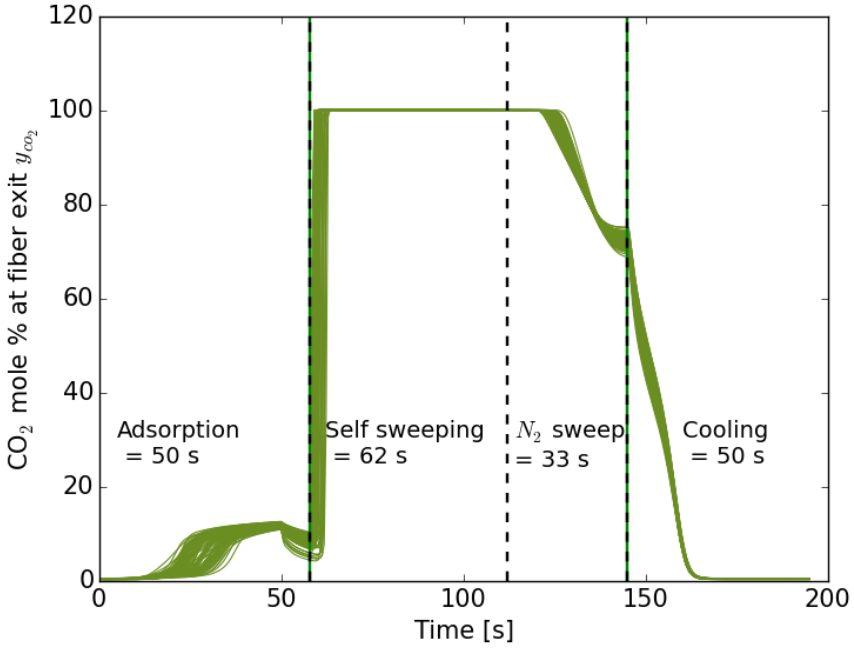


Figure 5.5: Varinflat-inf: Model prediction uncertainty in the exit CO_2 molar percent during the cycle at cyclic steady state (CSS)

approach (Fig.4.8) compared to "Hier-inf" approach (Fig.4.6), the predictive distributions also have a larger range of uncertainty compared to "Hier-inf". For example, the uncertainty in CO_2 swing capacity from "Varinflat-inf" approach is [0.36, 0.52] mmol/gfib (Fig.5.7) in comparison to [0.41, 0.52] mmol/gfib (Fig.5.3) in "Hier-inf" approach. The predictive distributions of CO_2 recovery and purity are shown in Fig.5.8(a) and Fig.5.8(b) respectively. As noted in the previous chapter, the disadvantage with this approach, however, is that the uncertainty in future predictive distribution is much larger, owing to the higher standard error assumption of C_{br} experimental data. Accordingly, in addition to the model prediction

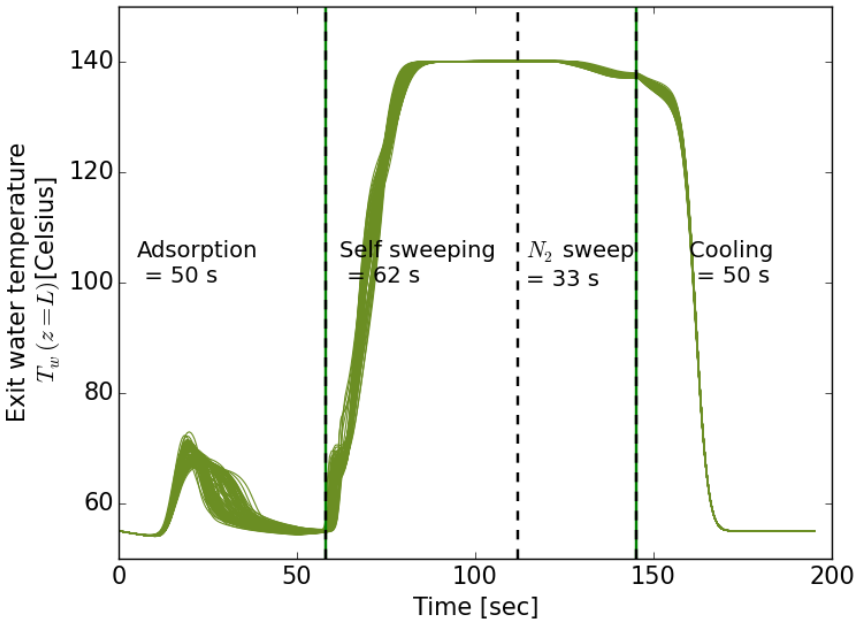


Figure 5.6: Varinflat-inf: Model prediction uncertainty in the exit water temperature T_w during the cycle at CSS

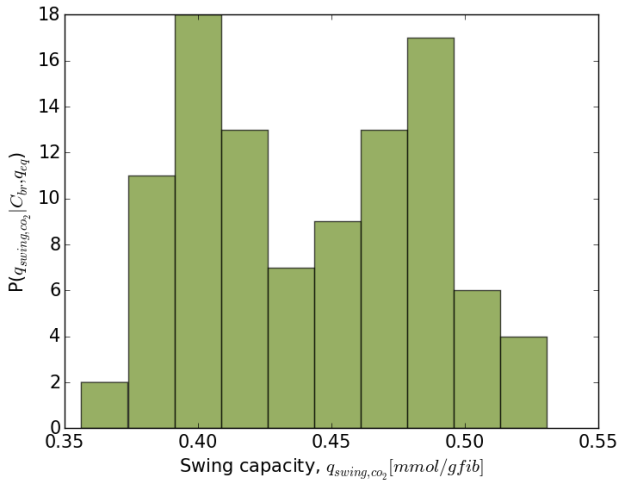
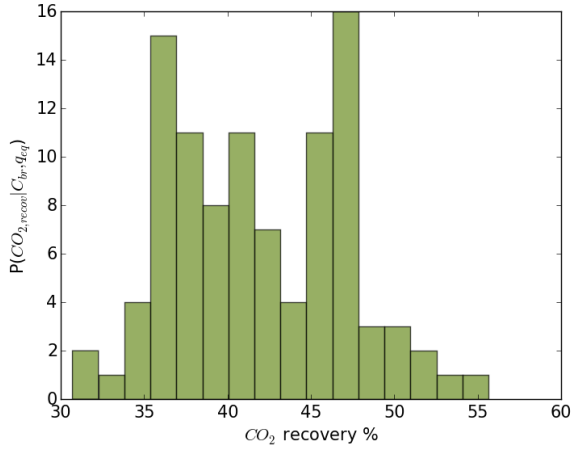
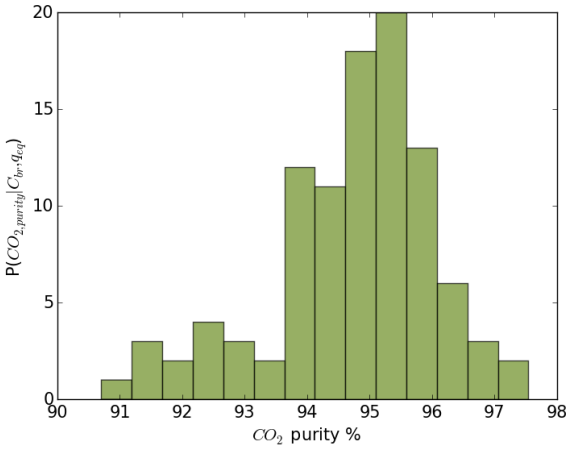


Figure 5.7: Varinflat-inf: Model prediction uncertainty in the CO_2 swing capacity at CSS



(a)



(b)

Figure 5.8: Varinflat-inf: Model prediction uncertainty in a) the CO_2 recovery and b) CO_2 purity at CSS

uncertainty observed in Figs.5.5-5.8(b), the future prediction will also include all the residual error components characterized during inference : $\tilde{y_{CO_2}} = \widehat{y_{CO_2}} + \epsilon_{var} + \epsilon_{noise}$. As the standard error of ϵ_{var} is relatively much larger compared to $\sigma_{C_g} = 0.25$ with $\sigma_{var} = 2.1$, the future predictions uncertainty would be much broader compared to "Hier-inf" approach.

5.2.3 Uresvar-inf - UQ summary

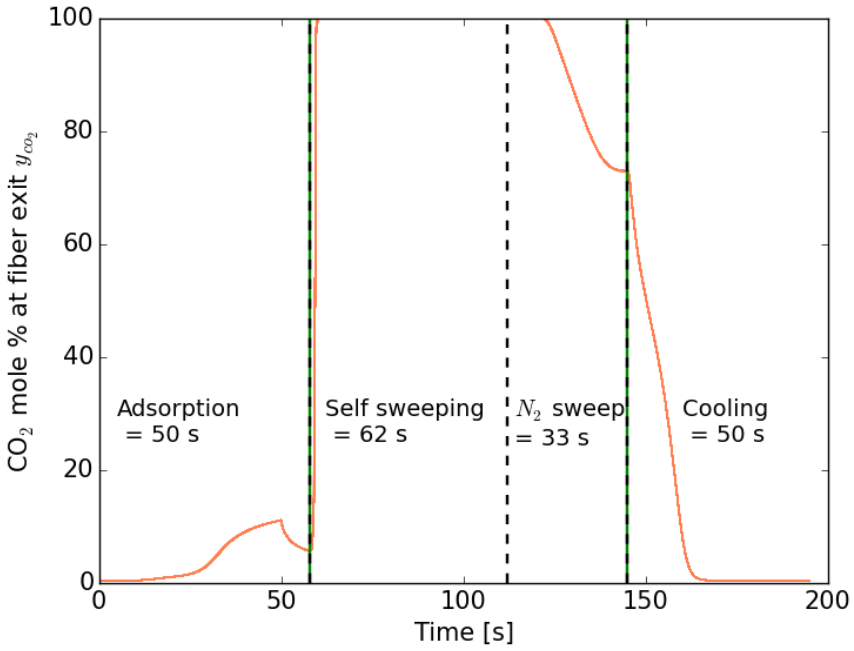


Figure 5.9: Uresvar-inf: Model prediction uncertainty in the exit CO_2 molar percent during the cycle at cyclic steady state (CSS)

As observed in previous chapter, the posterior parametric distribution that resulted from this approach, has a much reduced variance (Fig.4.12), as a result of incorporating the additional uncertainty in fiber packing within the module during characterization step. As the posterior parametric distributions are very narrow, the posterior predictive distributions obtained from propagation are also very narrow, giving almost a single value prediction for all the process variables as shown in Figs.5.9-5.12(b) for molar fraction of CO_2 at the outlet to the CO_2 purity respectively. Moreover, in contrary to the "Varinflat-inf" approach, the observation error in future prediction is also negligible due to lower standard

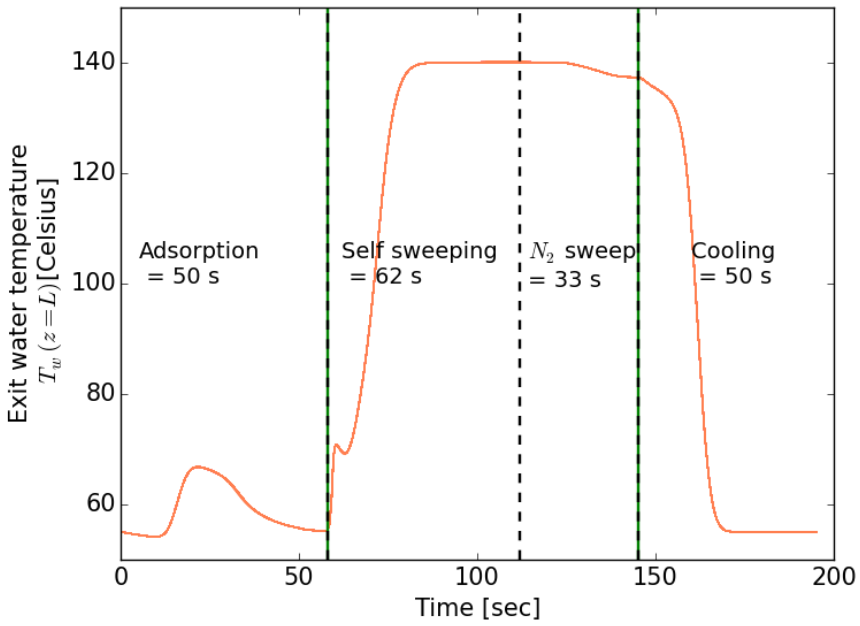


Figure 5.10: Uresvar-inf: Model prediction uncertainty in the exit water temperature T_w during the cycle at CSS

error, σ_{C_g} of measurement. However, as noted in chapter 4 under "Uresvar-inf" section, the posterior parametric distribution could not explain all the experimental data, especially those of 10-inch module within the error σ_{C_g}, due to the deterioration of particles and their corresponding diversity while sequentially updating the experimental data. Therefore, the prediction uncertainty obtained from this approach with a lower measurement error assumption cannot be considered as an accurate estimation of the complete uncertainty in the QMU evaluation. Perhaps, addition of model discrepancy term to account for the deviations in Fig.4.13(a), as in "Mdiscrep-inf" approach coupled with this approach is required.

5.2.4 Mdiscrep-inf - UQ summary

As the parametric distributions obtained under "Urevar-inf" approach could not explain all of the 10-inch data while still assuming small measurement error σ_{C_g}, as in "Uresvar-inf" approach, the model discrepancy term was added to the model prediction. The model

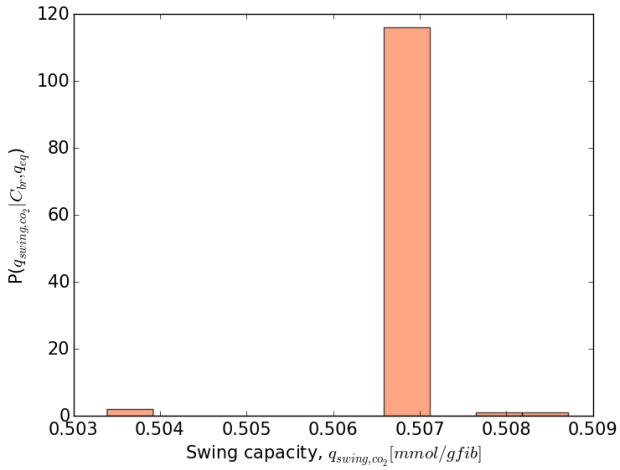


Figure 5.11: Uresvar-inf: Model prediction uncertainty in the CO_2 swing capacity at CSS

discrepancy term δ along with its hyper-parameters ϕ_t and σ_{model} were calibrated according to the CO_2 breakthrough experimental data (Fig.4.14, obtained during the adsorption step. Therefore, the calibrated model discrepancy is applicable only with respect to the prediction of CO_2 molar percent at the exit for larger module as described below in Eq.5.3.

$$C_{br} = \widehat{C_{br}} + \delta(\phi_t, \sigma_{model}) + \epsilon_{noise} \tag{5.3}$$

In order to estimate the contribution of model discrepancy on model variables, other than the one used during the calibration, such as water temperature T_w at the exit, CO_2 swing capacity, CO_2 purity and recovery, the respective model discrepancy terms, shown below, needs to have been calibrated.

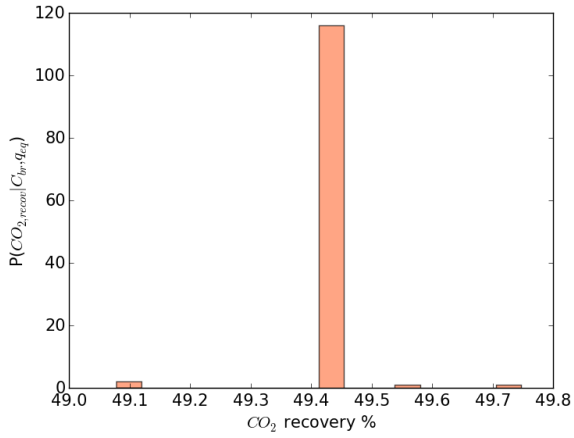
$$T_w = \widehat{T_w} + \delta_T(\phi_T, \sigma_{mT}) \tag{5.4}$$

$$q_{swing} = \widehat{q_{swing}} + \delta_{qs}(\phi_{qs}, \sigma_{mqs}) \tag{5.5}$$

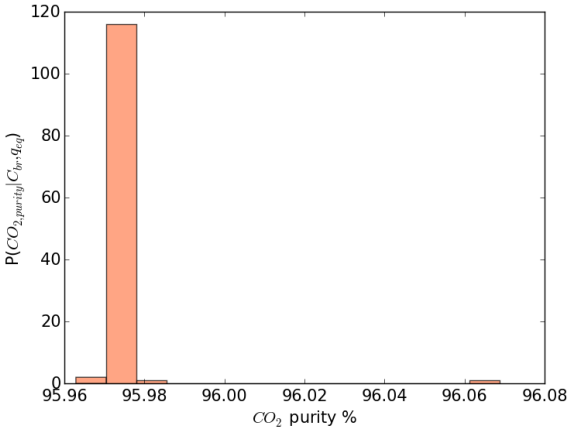
$$CO_{2,pur} = \widehat{CO_{2,pur}} + \delta_p(\phi_p, \sigma_{mp}) \tag{5.6}$$

$$CO_{2,rec} = \widehat{CO_{2,rec}} + \delta_r(\phi_r, \sigma_{mr}). \tag{5.7}$$

As the model is highly non-linear, it will also not be accurate to extrapolate the impact



(a)



(b)

Figure 5.12: Uresvar-inf: Model prediction uncertainty in a) the CO_2 recovery and b) CO_2 purity at CSS

of C_{br} model discrepancy on other process model variables. One bottleneck in determining the respective model discrepancy terms is that there are no experimental data available for the process variables T_w, CO_2 swing capacity, recovery and purity, which are needed during the calibration step. To overcome the difficulty, perhaps, synthetic data need to be used, which can be generated from the model using some guess values of parameters. These addition of appropriate model discrepancy terms to predict the impact on various process variables are again outside the scope of this thesis and can be pursued as a future research direction. Therefore, in this work, the model discrepancy is added only with respect to the prediction of CO_2 molar fraction at the exit as shown in Fig.5.13. As the model discrepancy

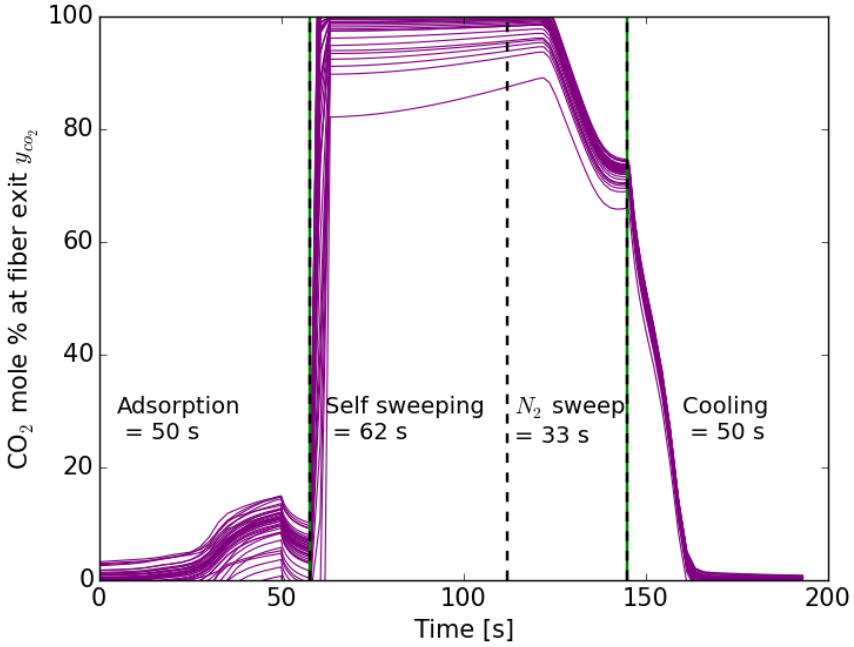


Figure 5.13: Mdiscrep-inf: Model prediction uncertainty in the exit CO_2 molar percent during the cycle at cyclic steady state (CSS)

δ, is a purely empirical term and is calibrated based on CO_2 molar composition during the adsorption step alone, its effect on the overall cycle including other steps might have been exaggerated as can be seen in Fig.5.13.

Thus, all the posterior parametric distributions obtained from all the four approaches used are propagated through the cyclic model to determine the predictive distributions of

Table 5.1: Comparison of prediction uncertainties with 99% probability among the four approaches

Approach	q_{swing} [mmol/gfib]	CO_2 purity [%]	CO_2 recovery [%]
"Hier-inf"	[0.41, 0.53]	[36, 55]	[91.5, 96.7]
"Varinflat-inf"	[0.36, 0.53]	[31, 55]	[91, 97.5]
"Uresvar-inf"	[0.503, 0.509]	[49.1, 49.7]	[95.96,96.06]
"Mdiscrep-inf"	N/A	N/A	N/A
"Conventional"	0.62	61	95

process variables of interest. A summary of the prediction uncertainty interval obtained from all the four approaches along with the values obtained from the conventional methods, is listed in Table.5.1. As seen from the table, the interval estimate from both the "Hier-inf" and "Varinflat-inf" are comparable and not very distinct, although the interval is wider in "Varinflat-inf" approach. However, the interval from "Uresvar-inf" is almost equivalent to a point estimate prediction and not comparable with the other two approaches, since the parametric distribution could not fit all the observed experimental data (Fig. 4.13). It can also be seen that the estimate obtained from conventional methods are very different from those of Bayesian methods. As some of the parameter values (for example A, D_{p0} and α) obtained from the conventional estimates (Table.3.3 and Table.3.1) are outside the tails of the posterior parametric distributions (Table.4.6) and therefore could be resulting in very different values outside of the interval of the prediction distributions.

This chapter summarized the UQ analysis of the process model, by propagating the uncertainties characterized based on the available experimental data, using various approaches that are studied in this thesis and yielded the QMU of the variables of interest. The next chapter, is on addressing the next question, which is to reduce the uncertainties in the model prediction variables once they are characterized and quantified.

CHAPTER 6

REDUCTION OF UNCERTAINTIES: OPTIMAL EXPERIMENTAL DESIGN

The uncertainties that are associated with model predictions may be unacceptable for some applications, which require a more definitive prediction from the model. In such cases, it is desirable to reduce the uncertainties by tackling them at their source. As, majority of uncertainties in model predictions arise due to errors in experimental data, both systemic and random, it is possible to reduce those uncertainties by gathering additional data that can reveal more information on the process behavior. However, it may often be difficult and expensive to perform experiments and collect additional data that are required to reduce the uncertainties. Under such scenarios, it is advantageous to determine the conditions at which performing experiments would provide maximum information gain. Such a strategy is termed as optimal experimental design (OED). By optimally designing the experiments, one can gain maximum amount of information about the system using the least number of experiments.

6.1 Bayesian design of experiments

In classical statistics, model based design of experiments is performed with the objective of reducing the variance either in the parameter estimates or in the model response (or prediction variables). Accordingly, there are several *alphabetic optimality* criterion such as A-optimality, D-optimality etc., which are used to define the objective function. Similarly, in Bayesian design of experiments, the objective function termed as utility function

is derived based on, whether the reduction of uncertainty is desired in parameters or in the model predictions. The advantage of Bayesian design, in comparison with the classical design, however, is that the classical design assumes normal distribution in likelihood and therefore in the distribution of parameters. Bayesian design, on the other hand, has no such assumption involved and therefore provides a more rigorous framework to reduce the involved uncertainties.

Most often, it is interesting and valuable to reduce the uncertainties in the prediction variables rather than in the model parameters. However, optimization towards reducing the uncertainties in prediction is computationally more expensive compared to optimizing experiments for parametric uncertainty reduction as it involves evaluating the model. On the other hand, reducing the uncertainties just in parameters alone need not proportionately translate to reduction of uncertainties in model predictions due to the varying levels of parametric sensitivity with respect to the model response variables. However, identifying the parameters for which the prediction variable have a higher sensitivity and reducing the uncertainties in those parameters could be a computationally tractable method to reduce the prediction uncertainties. This is the principle behind the optimal design algorithm that is developed and demonstrated in this chapter.

6.2 Problem formulation

In order to demonstrate the experimental design methodology that is proposed in this chapter (as discussed below), a simplified system of hollow fiber model considering uncertainties in fewer parameters is studied. The response variable whose uncertainty needs to be reduced is with respect to the CO_2 adsorption breakthrough capacity q_{br}, which is directly dependent on the CO_2 breakthrough curve and is defined as

$$q_{br} = q_{avg}(t)|_{t=t_{br}}, \tag{6.1}$$

where q_{qvg} is the total CO_2 adsorption capacity averaged along the length of the fiber and t_{br} is determined by the following equation.

$$t_{br} = \{\text{max } t | \, C_g(t, z = L) \leq C_{br}\} \tag{6.2}$$

The above equation determines the maximum time duration of the adsorption step such that the CO_2 concentration at the fiber exit (called as CO_2 breakthrough concentration) does not exceed the maximum concentration C_{brk}. In this work, the breakthrough concentration is considered as 10% of the inlet concentration $C_{brk} = 0.1C_{in}$. As the goal is to mainly demonstrate the proposed approach, some of the parameters, θ_{mass} are assumed to be at their deterministic values. Accordingly, the uncertainty in model prediction is assumed to be due to uncertainties in adsorption isotherm parameters alone. The experiments that are optimized are therefore, the TGA experiments that are used to collect the adsorption equilibrium capacity, q_{eq} which is directly related to all the adsorption isotherm parameters θ_{iso} (Eqns. 2.32 -2.35). The experimental design is aimed at determining the conditions, d (=$[T, P_{co_2}]$) at which the new data that is collected from the experiment, will reduce the uncertainty in prediction of q_{br}.

6.3 Utility function formulation

For experimental design optimization, the objective function, termed as *utility function*, needs to be a measure of uncertainty. In Bayesian statistics, it is defined using Shannon entropy, which is rooted in information theory. Shannon entropy is the negative logarithm of a probability distribution, whose negative is a measure of the information content of the distribution. In other words, maximum information content implies mininum uncertainty in a distribution. Accordingly, the relative information gain on observing a new experimental data at d can therefore be defined using the Kullback-Leibler (K-L) divergence criterion, which is the difference between the prior predictive and predictive distribution posterior to

observing the data at d. Defining $U(d)$ with regard to predictive distribution of q_{br}, we get

$$U(d) = \int_{q_{eq}} \int_{q_{br}} log \frac{P(q_{br}|q_{eq}, d)}{P(q_{br})} P(q_{br}|q_{eq}, d) P(q_{eq}|d) dq_{br} dq_{eq}. \tag{6.3}$$

Here, $P(q_{br}|q_{eq}, d)$ is the predictive distribution posterior to performing the experiment at d and $P(q_{br})$ refers to the predictive distribution prior to performing the experiment. Also, in the following equations, $Theta$ refers to the adsorption isotherm parameters θ_{iso}.

In the above formulation, the predictive distribution needs to be determined for several candidate design conditions, during which each evaluation by itself involves determining the respective parametric posterior distribution, followed by propagating the posterior parametric distribution through the model defined in Eqns.(2.1)-(2.35). Such an evaluation is computationally very expensive. The computational complexity involved can be significantly reduced by evaluating the utility function $U_p(d)$ as given in Eq.6.4, where the Shannon information criteria utilizing the K-L divergence is determined with respect to the parametric distribution rather than the predictive distribution as earlier (Eq.6.3), thereby eliminating the need to evaluate the predictive distribution.

$$U_p(d) = \int_{q_{eq}} \int_{\Theta} log \frac{P(\Theta|q_{eq}, d)}{P(\Theta)} \big| P(\Theta|q_{eq}, d) P(q_{eq}|d) d\Theta dq_{eq} \tag{6.4}$$

However, our objective is to determine the experimental condition that maximizes the information gain with respect to the prediction variable rather than the parameters themselves. Therefore, we determine and isolate the parameters which have a higher sensitivity for the prediction variable and evaluate the information gain with respect to the sensitive parameters alone, while holding the other parameters at their nominal values as defined in Eq.6.3.

$$U'(d) = \int_{q_{eq}} \int_{\Theta_s} log \frac{P(\Theta_S|q_{eq}, d)}{P(\Theta_S)} \big|_{\Theta_{NS}=\Theta_{mean}} P(\Theta|q_{eq}, d) P(q_{eq}|d) d\Theta dq_{eq} \tag{6.5}$$

where Θ_S refers to the selective subset of parameters from Θ to which the design variable

q_{br} is sensitive and Θ_{NS} refers to the remaining subset in Θ.This modified utility function as defined in can effectively provide a measure of the information gain with respect to the prediction uncertainty. The sensitivity of q_{br} with respect to adsorption isotherm parameters is evaluated using Global Sensitivity Analysis (GSA), which is explained in the subsequent section.

6.4 Global Sensitivity Analysis

Global Sensitivity Analysis (GSA) represents the sensitivity information with respect to the complete parametric space, in contrary to the local sensitivity analysis (LSA), which determine the sensitivity information at a local parametric value (Zhan *et al.* 2013). GSA is used especially when the parameters are correlated and have strong interactions with one another and/or vary over a large parameter space. Morris method, an one at a time (OAT) based screening method is used to determine the GSA of the adsorption isotherm parameters on the prediction variable (Campologno *et al.* 2007), (Iooss and Lemaitre, 2014).

In the OAT based Morris method, as the name suggests, each parameter is varied one at a time and the sensitivity of the respective parameter with respect to the output variable of interest is calculated. This local sensitivity is termed as an elementary effect (EE). Since the model output is a time dependent variable, a derived variable *viz.* the design variable q_{br} is used to determine the sensitivity values. In the notation used below, $\theta_{i,j}$ refers to ith sample of j^{th} parameter, in which the index j refers to the members in set θ and i refers to the row in the matrix domain Ω, which is the complete set of parametric values. The size of the matrix domain Ω is of the size ($row \times column$)=($p \times n$), which is representative of the joint parameter distribution. Here p refers to the size of the set θ and n refers to the number of samples which is representative of the joint parameter distribution. The elementary effect,

$EE_{i,j}$ for the parameter j with the i^{th} set of values is determined as follows.

$$EE_{i,j} = \frac{q_{br}(\theta_{i,1},..,\theta_{i,j-1},\theta_{\mathbf{i+1,j}},\theta_{i,j+1},..) - q_{br}(\theta_{i,1},..,\theta_{i,j-1},\theta_{\mathbf{i,j}},\theta_{i,j+1},..)}{\Delta\theta_j}\frac{\sigma_{\theta_j}}{\sigma_{q_{br}}} \quad i = 1...p; j = 1...n \tag{6.6}$$

In the above equation $EE_{.,j}$ is scaled using the standard deviation of q_{br} and θ_j for the numerator and denominator respectively (Sumner *et al.* 2012). To determine $EE_{i,j+1}$, all the θ_j are held at the previous values and only the $j+1^{th}$ parameter is incremented to $i+1^{th}$ value.

In this way, $EE_{i,.}$ of the i^{th} Morris trajectory are determined for all the parameters $j = 1..p$ and typically about r = 20 to 50 trajectories are evaluated, which are chosen randomly from the domain Ω. The sensitivity information of the parameters is quantified using two measures, which are μ_j and σ_j and are determined for every j^{th} parameter. They are determined from the individual elementary effects as follows.

$$\mu_j = \sum_{i=1}^{r} \frac{abs(EE_{i,j})}{r} \tag{6.7}$$

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^{r}\left(EE_{i,j} - \mu_j\right)^2}{r-1}} \tag{6.8}$$

Accordingly, the GSA of q_{br} with respect to all the adsorption isotherm parameters are determined by employing around $r = 50$ trajectories and reported in Table 6.1. While a high value of μ_j indicates a high sensitivity of q_{br} with respect to that parameter j, a higher value of σ_j is indicative of the nonlinear impact or a strong interaction of that parameter with other parameters. Accordingly, based on the relative σ_j values, A is found to have a strong nonlinear impact or a strong interaction with other parameters. And, based on the relative values of μ_j, it can be found that q_{br} is highly sensitive with respect to the parameters q_{m0}, A and lnb_0 having a high μ_j value and are included in the set θ_S. Whereas,

Table 6.1: Global sensitivity values of adsorption isotherm parameters

Parameters	**GSA mean** μ_j	**GSA** σ_j
q_{m0}	2.07	0.33
η	0.09	0.03
A	6.04	4.17
B	0.42	0.37
lnb_0	2.03	0.90
$-\Delta H_0$	0.05	0.04

the rest of the three parameters have a relatively low sensitivity and are included in θ_{NS}.The evaluation of GSA is performed in parallel with each of trajectory run in an individual processor and therefore 50 processors are used to determine the GSA value. The total computation time was nearly 9 hours of CPU time, as some of the combinations of isotherm parameters make the time integration in Eqns.2.1- 2.16, stiff and time consuming.

With the parameters being classified based on their sensitivity, the next step is to determine the optimal design condition, d_{opt} that maximizes the modified utility function $U'(d)$ and is formulated as :

$$d_{opt} = \arg\max_{\mathrm{d}\in\mathcal{D}} \mathrm{U}'(\mathrm{d}), \tag{6.9}$$

where $\mathcal{D}$ is the design space of the two variables $[T, P_{co_2}]$.

6.5 Results of Sequential OED

Optimal experimental design (OED) can be determined either simultaneously or sequentially. In the simultaneous design, all the experimental design conditions are determined at the same time and chooses the set of conditions that maximizes the overall gain in information with all the data considered together. The sequential design, on the other hand, determines one optimal design at a time and is based on the incremental gain of information that the next experiment would provide over what is available from the existing set of experiments. Sequential design may be advantageous over the simultaneous design, as it incrementally updates the information on the system behavior during its subsequent design.

This could be a significant aspect in cases, where the system may behave differently in different regions of the design space. For example, in a sequential design, it is possible to be more accurate in the estimation of heteroscedastic variance measurement error variance, which cannot be known apriori in case of simultaneous design approach.

6.5.1 Initial uncertainty in sequential design

Optimal design in this thesis is determined based on the approach of sequential design, wherein the initial design set, d^o, is assumed to be non-empty with a set of experimental design conditions as shown in Fig.6.1. The adsorption equilibrium capacity measurements are those for *Class 2* amine sorbents. On the other hand, the data that has been used so far are those related to that of *Class 1* amine sorbents. Although, the sorbents are different, the same process model described earlier is applicable to the *Class 2* sorbents based hollow fiber as well. However, few minor simplifications are made to the overall mass transfer coefficient model as follows:

$$K_{ov,bulk} = K_b e^{-\alpha q_{bulk}} \tag{6.10}$$

and $K_{ov,surf}$ is fixed at a constant value. The values of the all the process parameter values relevant to the *Class 2* sorbents based hollow fiber module is shown in Table.6.2.

To begin with, the base case (prior to performing OED) uncertainty in the adsorption isotherm parameters is determined by performing MCMC simulation with the data shown in Fig.6.1 and the obtained posterior parametric distributions are shown in Fig. 6.2. The posterior parametric distribution, thus obtained is propagated through the hollow fiber model 2.1-2.35, to predict the distributions of CO_2 breakthrough concentration, C_{br} and the corresponding adsorption breakthrough capacity q_{br} and are shown in Figs.6.3. The adsorption step simulation is performed at a fluegas flow rate of 40 mL/min and T_{mod} = 35^o C.

Table 6.2: Values of parameters and properties (deterministic parameters)

Parameters	**Values**	**Units**
A_t	6.40×10^{-7}	m^2
C_{pf}	1200	J/kgK
C_{pg}	1094	J/kgK
C_{pt}	500	J/kgK
C_{g0}	0.00	mol/m^3
$C_{g,in}$	5.40	mol/m^3
h_g	40.0	W/m^2K
h_t	2500	W/m^2K
$K_{ov,surface}$	1.00	s^{-1}
$K_{ov,b}$	0.30	s^{-1}
L	0.25	m
L_t	0.05	m
r_i	225	μm
r_o	375	μm
r_{fs}	809	μm
u_g	0.07	m/s
U	60.0	W/m^2K
α	5.1	$g.fiber/mmol$
ϵ_f	0.48	
ΔH_{ads}	-65.0	kJ/mol
ψ	0.11	
ρ_f	960	kg/m^3
ρ_g	1.20	kg/m^3

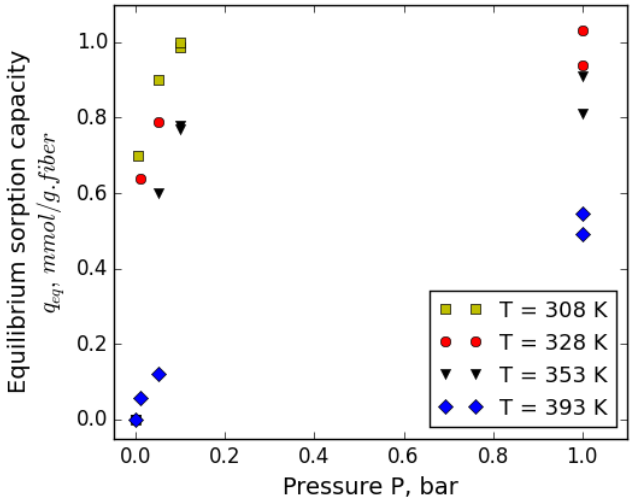
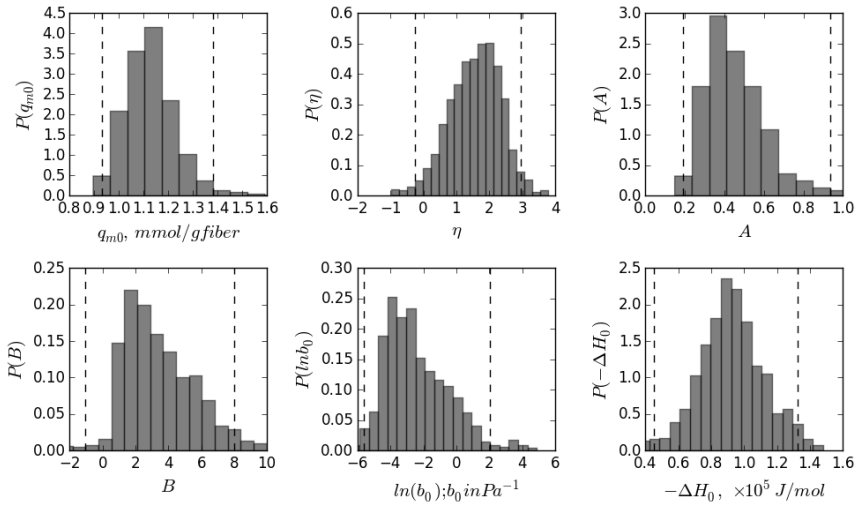
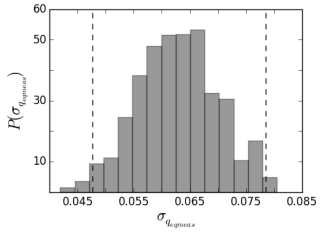


Figure 6.1: Initial experimental design set, d^o used in sequential optimal design along with values of observable, the adsorption equilibrium capacity q_{eq} at $[d^o \in [T, P_{co_2}]$

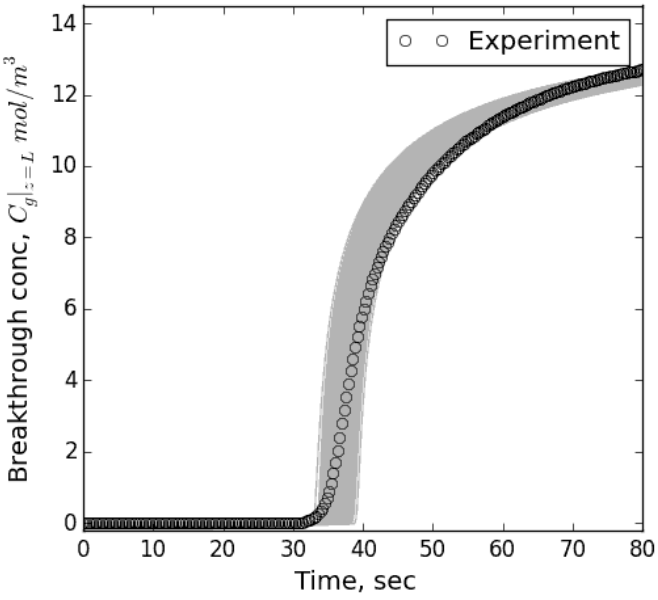


(a)

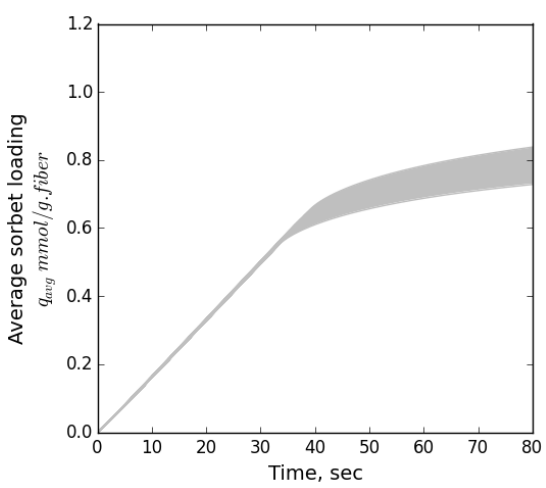


(b)

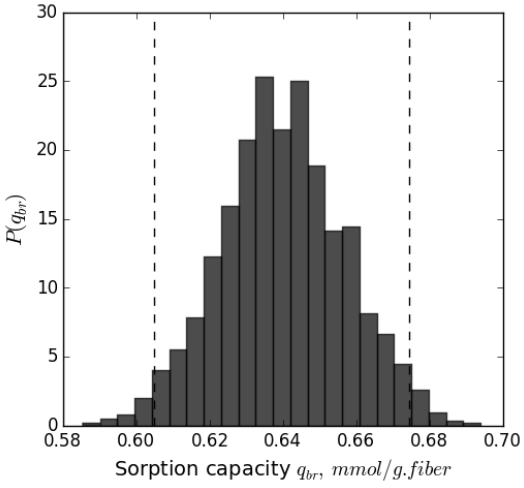
Figure 6.2: Posterior parametric distribution, $P(\Theta|q_{eq})$ of a) of the adsorption isotherm parameters and b) hyperparameter, for the base case with initial data prior to OED; 95% credible interval shown by the dashed lines.



(a)



(b)



(c)

Figure 6.3: Posterior predictive distributions of base case for a) $CO2$ breakthrough concentrations and, b)average sorption capacity q_{avg} over the adsorption time and c) distribution of breakthrough capacity q_{br}, with dashed line indicating the 95% credible interval.

6.5.2 Utility function evaluation

With the base case established, the search for the new optimal experimental design point is performed via a discretized grid search approach, instead of a continuous search over the $\mathcal{D}$ design space. Optimization of the utility function $U'(d)$, therefore involves determining the posterior parametric distribution at each of the candidate design point in the grid space $\mathcal{D}$. Hence, each of the evaluation is still computationally expensive, regardless of the modification in the formulation of the actual utility function $U(d)$, which is with respect to predictive distributions, to be in terms of posterior parametric distribution as $U'(d)$. As evaluation of $U'(d)$ at each grid point is independent of the other, each enumeration is performed in parallel across multiple processors. The design point with the highest value of $U'(d)$, is then chosen as the optimal condition d_{opt} to perform the new experiment.

As the utility function in Eq.6.3 does not have a closed form (i.e an analytical solution), it is evaluated using Monte Carlo estimation as follows.

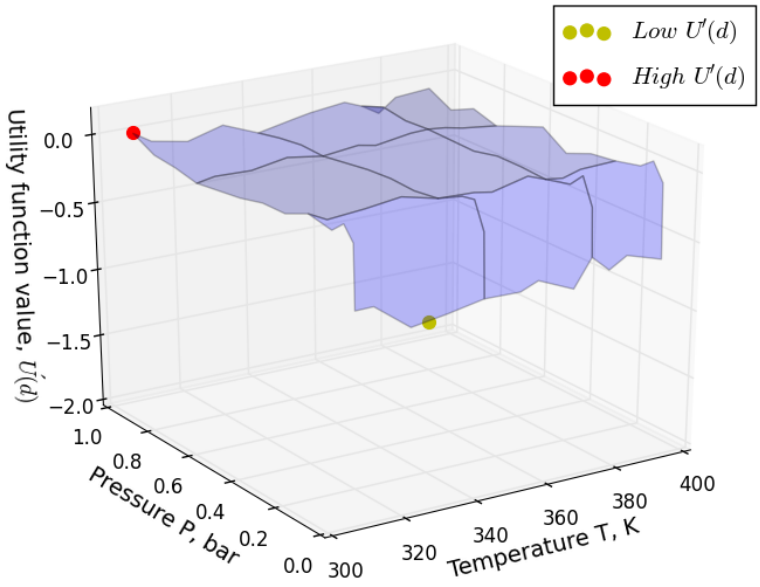
$$U'(d) \approx \frac{1}{n_{obs}} \frac{1}{n_{mcmc}} \sum_{i=1}^{n_{mcmc}} \sum_{j=1}^{n_{obs}} \{logP(q_{eq,j}|\Theta_{S,i}, d) - logP(q_{eq,j}|d)\}\big|_{\theta_{NS,i}=\theta_{mean}}. \quad (6.11)$$

In the above equation, the second term is evaluated as follows:

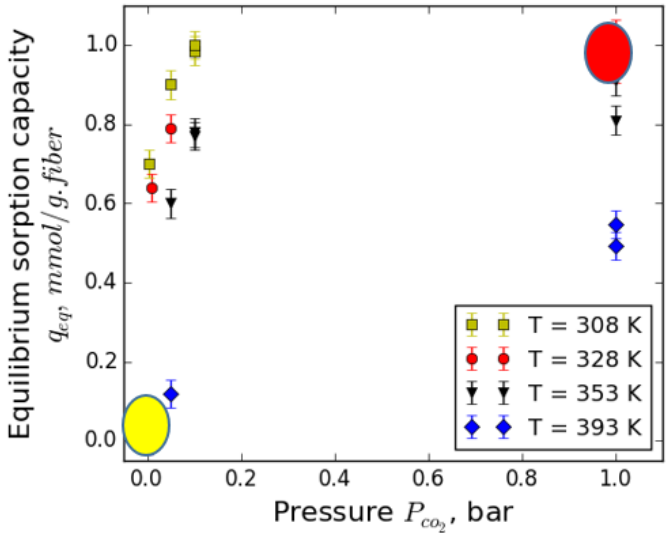
$$P(q_{eq,j}|d) = \int_{\Theta} P(q_{eq,j}|\Theta_S, d)|_{\Theta_{NS}=\Theta_{mean}} P(\Theta|q_{eq}) d\Theta \quad (6.12)$$

$$\approx \frac{1}{n_{mcmc}} \sum_{i=1}^{n_{mcmc}} P(q_{eq,j}|\Theta_{S,i})|_{\Theta_{NS,i}=\Theta_{mean}}. \quad (6.13)$$

The 2D grid of $d \in (T, P)$ is designed with 144 grid points as $T \times P = 12 \times 12$, which spans over the range of required temperatures and CO_2 partial pressures and is plotted in Fig.6.4(a) along with the values of the utility function $U'(d)$ obtained at the respective grid point. The evaluations are performed in 72 processors with two grid point evaluation per processor and total CPU time was nearly 40 mins.



(a)



(b)

Figure 6.4: a) Utility function surface $U'(d)$ over the design space of $[T, P]$ with two design points chosen in the regions of low and high utility value respectively b) regions of chosen high and utility design points are marked in the adsorption equilibrium data plot

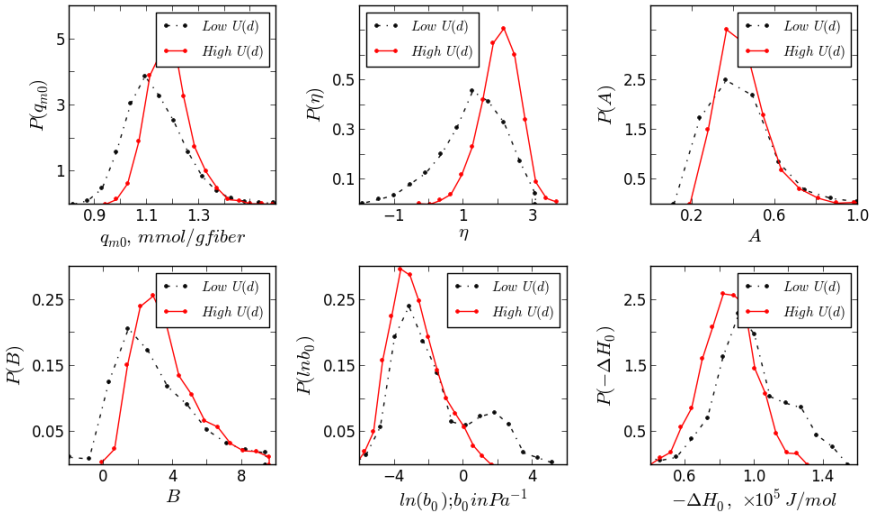
6.5.3 Choice of optimal design condition d_{opt}

From the utility function surface $U'(d)$ plotted in Fig.6.4, it can be seen that except the design condition at very low pressures around 0.005 bar, all other pressure and temperature values seem to have nearly same amount of potential information gain. This can be explained due to the reason of constant variance, $\sigma_{q_{eq}}$, assumption of measurement error. Consequently, an experimental data at low pressures of 0.005 bar can have a very minimal information gain due to low value of q_{eq} at low pressures (Fig.6.1 and Eq.2.32) with a relatively larger error. On the other hand, at higher pressures, q_{eq} has a relatively larger value and can have a potential information gain because of the smaller measurement error.

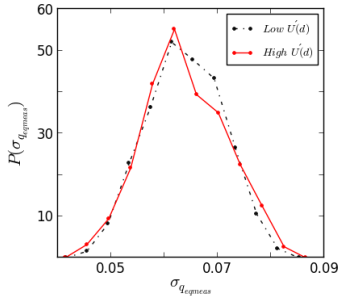
To verify and validate the relative information gain at different design points as indicated by the utility function surface, two cases are studied each of which adds an experimental data q_{eq} at low utility and high utility values $U'(d)$ respectively. A design point with a high utility function value is chosen at (35^{o} C, 1 bar) and an other case with a low utility function value is chosen at (55^{o} C, 0.005 bar). The utility functions of the chosen points are depicted in Fig. 6.4(a), with their corresponding regions in the equilibrium data plot depicted in Fig.6.4(b). TGA experiments are performed at both the design conditions (55^{o} C, 0.005 bar) and (35^{o} C, 1 bar) and the q_{eq} values are found as 0.56 $mmol/gfib$ and 0.91 $mmol/gfib$ respectively. There are now two cases of data sets, one with the new data at (55^{o} C, 0.005 bar) added to the original data set and other case with the new data at (35^{o} C, 1 bar) added to the original data set. The former case is labelled as 'low $U'(d)$' and the latter one is labelled as 'high $U'(d)$'. The original data set is indicated in Fig.6.1.

MCMC simulations are performed for both the cases with their respective data set. The posterior distribution of the isotherm parameters are obtained from both cases are compared against one another in Fig.6.5. The 95% credible intervals (CI) are compared for the two cases for all the parameters in Table.6.3. As seen from the table, the case where a data point is added at higher $U'(d)$ has a smaller uncertainty or narrower CI for almost all the adsorption isotherm parameters compared to the case where a data point is added

at a lower utility value. The hyperparameter $\sigma_{q_{eq}}$, however, shows no significant difference in the uncertainty range, which indicates that the measurement error by itself has not been reduced, partly due to the fact that we assume a constant standard deviation of measurement error. However, the actual goal of the OED is to reduce the uncertainty in q_{br} rather than in parameters themselves.



(a)



(b)

Figure 6.5: Comparison of the a) isotherm posterior distribution of parameters $P(\Theta|q_{eq}, d)$ and of the b) hyperparameter posterior distribution, between the two cases of adding a data corresponding to high and low $U'(d)$ respectively

Therefore, using the posterior parametric distribution obtained for the two cases, the respective predictive distributions for the CO_2 breakthrough concentration profiles and q_{br} are determined and compared against one another in Figs.6.6 and Fig.6.7 respectively. From

Table 6.3: Comparison of credible intervals between the two cases of $U(d)$ for all the parameters

Parameters	**95$\%$ CI for high** $U'(d)$	**95$\%$ CI for low** $U'(d)$
q_{m0}	[1.03,1.36]	[0.93,1.38]
η	[0.75, 3.09]	[-0.25,2.95]
A	[0.3,0.73]	[0.19,0.93]
B	[0.75,8.5]	[-1.04,7.98]
lnb_0	[-5.30,0.24]	[-5.59,2.05]
$-\Delta H_0$	[0.55, 1.5]	[0.45,1.32]
$\sigma_{q_{eq}}$	[0.05,0.08]	[0.05,0.08]

the Figure.6.6, it can be clearly seen that the case with the addition of an experiment having a higher utility value clearly has a lower uncertainty in breakthrough profile distribution. Accordingly, the uncertainty in the breakthrough capacity, q_{br}, a derived process variable, also is much lower in 'high $U'(d)$' case compared to 'low $U(d)$' case. Based on the 95$\%$ credible intervals shown of $P(q_{br})$, the maximum sorption capacity that can be obtained without violating the breakthrough concentration constraint of CO_2 at the exit is increased from the value of 0.605 $mmol/g.fiber$ in the base case (Fig.6.3(d)) to 0.63 $mmol/g.fiber$ (Fig.6.6). The increase is around 5$\%$ of the original capacity, with an addition of a single experimental point having a higher utility value. On the other hand, the addition of an experimental point having a lower utility value did not make any significant difference in the uncertainty compared to the base case. To give an economic perspective of the gain due to uncertainty reduction, Kulkarni and Sholl showed that an increase in the adsorption breakthrough capacity by 25$\%$ can result in the reduction of the net cost of CO_2 capture by 23$\%$(reduction from approximately 110\$/tCO_2 to 90\$/tCO_2).

The approach that is propsed in this thesis is successfully demonstrated to reduce prediction uncertainty without having to optimize the original utility function involving the computationally intensive predictive posterior distributions. The sequential design procedure can be iteratively repeated to gather new data until the final uncertainty reaches the acceptable limit or a non-reducible error limit, i.e when the model mismatch errors cannot be reduced further without making any changes to the model structure etc or incorporating

additional uncertainty in other existing model parameters. The next chapter summarizes the uncertainty framework that has been developed in this thesis and provides some concluding remarks and future research directions in continuation of this thesis.

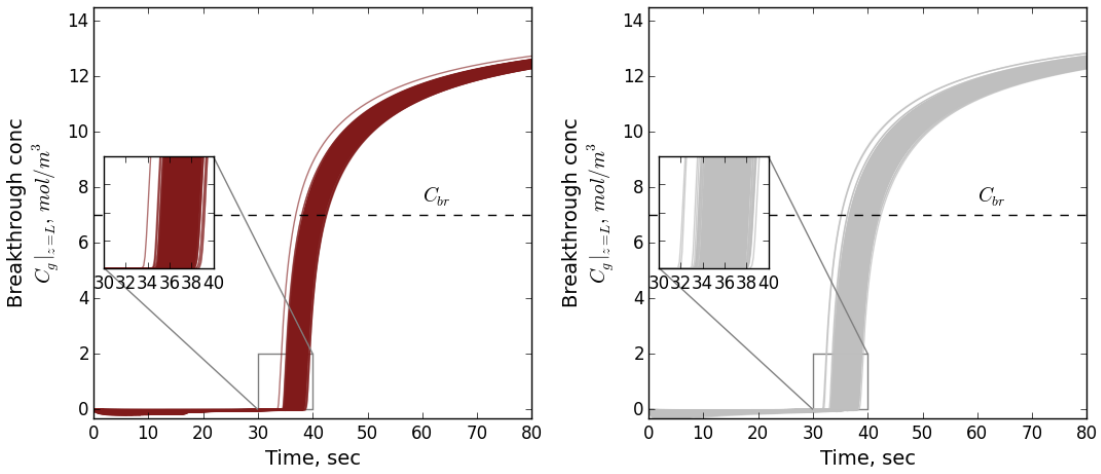


Figure 6.6: Comparison of the breakthrough profile distribution between the two cases of added experimental data at different utility $U'(d)$ values, with a)an experiment at high $U'(d)$ added b) an experiment at low $U'(d)$ added

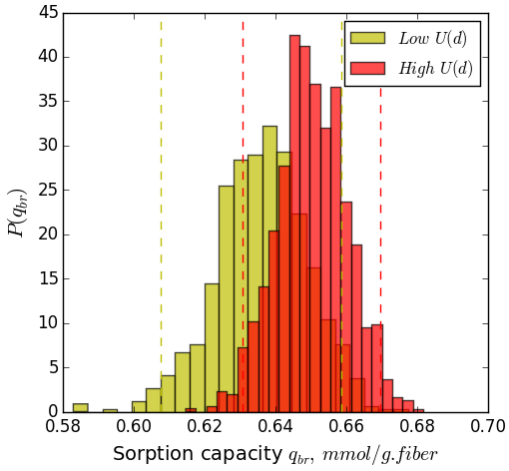


Figure 6.7: Comparison of the predictive distribution between the two cases with the corresponding dashed lines bounding 95% CI

CHAPTER 7

CONCLUSIONS, CONTRIBUTIONS AND FUTURE RESEARCH DIRECTIONS

In this thesis, a comprehensive Bayesian framework of methodology to perform uncertainty analysis on complex physical process model predictions was proposed and demonstrated for an application study of post-combustion carbon capture using Rapid Thermal Swing Adsorption (RTSA) process. The framework includes all the three essential components of uncertainty analysis, viz: characterization of uncertainties, followed by propagation to determine their impact on model predictions and finally designing optimal experiments to reduce uncertainties in model prediction as shown in Fig.7.1. The carbon-di-oxide sepa-

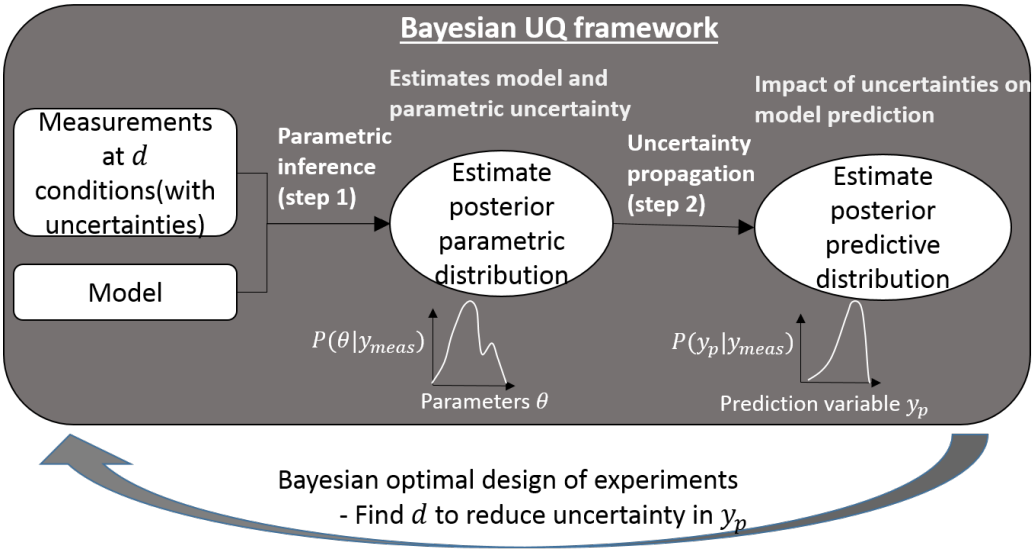


Figure 7.1: UQ framework under Bayesian analysis

ration by RTSA in a hollow fiber sorbent is a dynamic process with a complex interplay of non-isothermal mass transfer and equilibrium effects. The detailed first principles based

mathematical model of the process developed in this thesis with minimum simplifying assumptions around mass transfer resistance, however, was able to accurately simulate the process behavior and predict the asymmetric CO_2 breakthrough curve, with a sharp breakthrough and a long asymptotic tail. The process model was developed in gPROMS, a commercial process modeling software [58]. All the physical model parameters involved in the process, viz: adsorption isotherm parameters, mass and heat transfer parameters, were initially estimated through conventional methods viz : Maximum Likelihood Estimation (MLE) and Ordinary Least Squares (OLS) regression. Two different experimental data were considered to estimate the parameters. The measurements of adsorption equilibrium capacity q_{eq} from Thermo Gravimetric Analysis (TGA) experiments were used to estimate adsorption isotherm parameters. Mass transfer parameters, on the other hand, were estimated using measurements of CO_2 breakthrough measurements at different temperatures and flue gas flow rates.

Although, the model with the estimated parameters could predict the behavior very well at certain operating conditions (Figs.3.5(a)), the deviation in model prediction from the experiment was large at different conditions (Fig.3.5(b) and Table.3.4). The reason for not being able to predict accurately the process behavior at various conditions, regardless of a detailed process model, was the presence of various errors and inaccuracies (both random and epistemic) both in the experimental data and in the model. Therefore, in order for the model to be able predict the process behavior at various operating conditions, it was found to be necessary to incorporate those errors and inaccuracies as uncertainties in the model. Thereby, the mathematical model predictions, will include the effect of those uncertainties and be applicable for a range of operating conditions.

Bayesian approach was chosen to perform the uncertainty analysis as it offers several advantages, viz: inclusion of prior information, sequential updation of data via SBU and providing unifying framework to perform the complete uncertainty analysis. However, performing Bayesian inference for a complex physio-chemical process such as RTSA, es-

pecially without model reduction, is computationally challenging and has never been studied before in literature. The computational complexity of performing Bayesian parametric inference was handled by use of Sequential Monte Carlo (SMC), a parallel simulation algorithm which is based on the principle of particle tracking algorithm. The entire implementation of SMC is done in Python, from the scratch with parallelism using MPI framework. In addition, the entire process model was re-implemented in Python (with explicit spatial discretization schemes), in order to avoid cross-compiling issues with gPROMS that would be required while simulating the gPROMS process model during Bayesian inference. Four different approaches were studied to characterize the uncertainties that are involved in the process model parameters and Bayesian inference was performed in each case to determine the posterior distribution of the parameters involved.

The first approach, named as "Hier-inf", was based on separating the experimental data to different subsets and estimating the posterior distribution for each subset separately. The subsets of data were grouped based on the module that is used to collect the experimental data, as it was expected that there is a random variability associated with the synthesis of each fiber module. Accordingly, the breakthrough data are grouped as that of 10-inch module and 17-ich module. The resulting posterior parametric distributions obtained from the inference of the separated sub-sets of data were able to fit and explain only their respective data sets very well (Fig.4.7), within the measurement error σ_{C_g}. This approach resulted in two different separated distributions of parameters in their posteriors (Fig.4.6, differing even in some of the fundamental parameters (q_{m0}, $log(b_0)$, A, B and η) which are not expected to vary with the module synthesis errors. It could be due to the high correlation that exists between the parameters (Table. 4.4), which could have lead to the parameters identifying different modes while the highly correlated ones retain their distribution. Once the existing parametric uncertainties are characterized with their posterior parametric distributions, they are propagated through the model to determine the corresponding uncertainties on the model predictions. The prediction uncertainties with respect

Table 7.1: Comparison of future prediction uncertainties among the four approaches

Approach	Residual standard error (RSE) of ϵ_y	Future observation $\tilde{y}$
"Hier-inf"	0.25	$\widehat{y} + \epsilon_y$
"Varinflat-inf"	2.1	$\widehat{y} + \epsilon_y$
"Uresvar-inf"	0.21	$\widehat{y} + \epsilon_y$
"Mdiscrep-inf"	0.21	$\widehat{y} + \epsilon_y + \delta_y$

to the CO_2 mole composition at the exit, exit water temperature, average swing capacity, and CO_2 purity and recovery are determined. As the posterior parameters are obtained as two separated distributions in "Hier-inf" approach, the propagation also results in two separated distributions for all the prediction variables. This implies that a choice needs to be made on whether the outcome from 10-inch module needs to be believed or the one based on propagation of posterior distribution from 17-inch module. Accordingly, based on the choice of belief, one could get multiple prediction estimates accordingly. In this work, equal probability is considered for both the distributions and the analysis are made based on that assumption. A comparison of the prediction uncertainty of a future observation from all the four approaches is shown in Table.7.1. As the variance of residuals is very low in this approach, the resulting prediction uncertainty of a future observation in this approach is low Table.7.1), and includes mostly the uncertainties of parameters alone. The second approach, "Varinflat-inf", in contrast to "Hier-inf" resulted in a unique broader range (Figs.4.8) of parametric distributions. This was due to the assumption of larger residual errors, as a result of lumping all the "residual variability" in the Gaussian uncorrelated error term. The prediction uncertainty of the performance variables CO_2 recovery, purity and the other variables, by itself, was not significantly broader compared to those from "Hier-inf" approach (Figs.5.5 -5.8(a) and Table. 5.1). The main disadvantage of this approach, however, is that prediction uncertainty of a future observation, which includes all the residual error, is enlarged in this approach (Table. 7.1) and might have a much larger and disproportionate impact on the variables other than the observed ones (ie. CO_2 breakthrough curve) used in parametric inference.

The third approach, "Uresvar-inf" was attempted to reduce the residual errors (as in the previous approach), while simultaneously attempting to determine a unique distribution. It does so, by unfolding the "residual variability" though consideration of uncertainties in additional parameters of the model. As the fiber packing within the module is not controlled to be uniform in the experimental module, the free cross section around each fiber can vary significantly from one to another, ultimately varying the module outlet CO_2 breakthrough concentration. The parameter r_{fs}, the Happel's free surface radius is thus assumed to be random within the module and the uncertainty in r_{fs} of each of the six fibers within both the modules are characterized in addition to the model parameters, while performing the Bayesian inference. The variance of the resulting posterior parametric distribution is drastically reduced (Table.4.6), in comparison with the previous two approaches. Although, a unique distribution was obtained, the parametric posterior could not fit all the data within the measurement error. This could be due to the particle depletion during the Sequential Bayesian Updating (SBU) that is performed using Gaussian KDE of prior (posterior with old data). As a result, the diversity and the variance of the distribution are lost during the prior updation with the some of the intermediate experimental data, thereby leading to the loss of information of the old data. As the parametric distribution obtained from this approach could not fit all the observed experimental data, the prediction uncertainty obtained from this approach is not directly comparable with those of the two approaches.

As it was found to be quite difficult to determine a unique parametric distribution within the measurement error using the existing model structure, regardless of the addition of uncertainties to the model parameters, it lead to the conclusion that the model might be inadequate to explain all the data. Therefore, in order to account for the deviation between the model prediction and the experimental data of "Uresvar-inf" approach, additional uncertainty was introduced in the model. This is termed as the "Mdiscrep-inf" approach, in which the dynamic model discrepancy is characterized based on the deviation that was observed in "Uresvar-inf" approach between the model prediction and experimental data.

This approach introduces additional hyper-parameters to characterize the model discrepancy term (Table.4.7), which could result in identifiability issue that occurs when there are more parameters to be inferred with the available data. However, in the case of the breakthrough data, the addition of model discrepancy to the "Uresvar-inf" approach was indeed able to explain all the observed experimental data (Fig.4.15). The disadvantage with "Mdiscrep-inf", however, is that the model discrepancy, being an empirical term, needs to be calibrated for each of the model variables of interest. As the model discrepancy was calibrated only with respect to the $\mathrm{CO_2}$ outlet mole fraction in this thesis, it was not possible to determine the contribution of model discrepancy to other prediction variables, viz: $\mathrm{CO_2}$ recovery, purity, CO_2 swing capacity. Therefore, it was not possible to compare the prediction uncertainty of "Uresvar-inf" approach against the other two approaches (Table. 5.1). The prediction uncertainty involved for a future observation, in this case, includes the model discrepancy in addition to the measurement noise error (Table.7.1). On comparing all the approaches that are studied in this thesis, "Hier-inf" seems like a good approach to characterize the uncertainties involved in a complex process model, by grouping the data into series and inferring them separately. However, as some of the fundamental parameters, which were not expected to vary with the change of model, had distinct distributions from the two modules's data, the posterior distributions hence obtained does not seem like a physically sensible distribution. Considering that, the alternative "Varinflat-inf" seems like the best approach as it provided a unique posterior parametric distribution, although the resulting prediction uncertainty in future observation is much larger in this approach.

The last component of uncertainty analysis, the optimal design of experiments to reduce the uncertainties in prediction variables further, is demonstrated for a simpler system with a few uncertain parameters. Sequential design is considered while optimizing experiments, where experiments are added one by one based on the results that are obtained so far. The utility function, objective function defining the value of incremental information gain is defined with respect to the reduction in prediction uncertainty of adsorption

breakthrough capacity q_{br}. However, to reduce the computational overhead involved in performing the prediction for every candidate, the utility function is reformulated in terms of posterior parametric distribution, considering only with respect to those parameters for the adsorption breakthrough capacity has a high sensitivity. Thereby, the reduction in parametric distribution uncertainty can directly translate to reduction in prediction uncertainty. The proposed algorithm is demonstrated with the addition of a single experimental data, chosen from two different regions of utility function, one at a lower value $U'(d)$ and other at a higher value of $U'(d)$ (Fig.6.4). Based on the results of prediction distribution of q_{br} (in Fig.6.7), it is clearly shown that by targeting a reduction in posterior parametric distribution with respect to the sensitive parameters alone, it is possible to reduce the prediction uncertainty.

Thus, a comprehensive Bayesian framework of uncertainty analysis was developed, which can applied for any complex chemical process with the use of the appropriate process model in place of the RTSA model used. Hence, this thesis advances the state of art in the uncertainty analysis of complex chemical processes.

7.1 Future research directions

This thesis targeted at answering the larger question of how to perform uncertainty analysis for a complex chemical process in a computationally feasible manner, while retaining the complexity of the model without resorting to any model reduction. The results obtained for that problem lead to further interesting questions, some of which could be pursued as future research in this area.

7.1.1 Coupling surrogate model with SMC algorithm

This thesis was aimed at utilizing the complete model, without resorting to order reduction while characterizing the uncertainties. Consequently, for a complex process, Bayesian inference gets computationally expensive, regardless of parallelism using SMC algorithm,

taking atleast upto about 15 days (Table. 4.7) to obtain the posterior parametric distribution. Although it is more accurate and effective to use the complete model for uncertainty analysis, especially for models with non linearly dependent parameters, for which determining the surrogate is a huge challenge [27]. However, for processes, where the surrogate model development is feasible, coupling SMC algorithm with surrogate modeling will provide a significant speedup in obtaining the posterior distribution. In particular, employing a polynomial chaos expansion method such as stochastic Galerkin can reduce the computational effort involved in a single simulation tremendously.

7.1.2 Rigorous model discrepancy approach

The model discrepancy "Mdiscrep-inf" approach, has been applied only with respect to the observable, the CO_2 mole fraction at the exit due to the lack of experimental data with respect to the other observables of interest such as CO_2 purity, recovery and CO_2 swing capacity. In case of RTSA experiments, it is possible to measure some of these variables through performing cyclic experiments. However, in many processes, it may not be possible to obtain data from the experimental set-up. In such cases, using sythetic data generated based on the parameters obtained during characterization step may be an alternative. Another bottleneck with the extension of model discrepancy to many process variables, is the issue of identifiability, which occurs when hyper-parameters unrelated to process model needs to be calibrated[68]. Developing a framework, that can extend the model discrepancy approach to prediction variables by addressing the issue of identifiability is another possible research directions in the future.

7.1.3 Decomposition of model to conditionally independent sub models

In this thesis, the complete RTSA process model was decomposed to sub-models of adsorption isotherm equilibrium (Eq.2.32-Eq.2.35) and separate experiments were designed to collect the adsorption equilibrium capacity data. The data, thus obtained were used to

directly estimate or perform Bayesian inference for the isotherm parameters separately, isolating from the rest of the model complexity. Along the same principle, a complex model involving multiple unknown parameters can be decomposed to relatively simpler models with fewer parameters. Thereby, separate experiments can be designed to collect the appropriate data relevant to the identified simpler models, which can be used subsequently in the estimation of the corresponding parameters. Performing such inferences of simpler models can drastically reduce the problem of "dimensionality" that occurs when large number of parameters have to be inferred in complex systems [37]. In case of adsorption isotherm models, it was quite direct to identify the model due to its isolated and separable presence in the entire model. There may still be other simpler models relating to different parameters and process variables that can be obtained by rigorously decomposing the complete model. Developing an algorithm that can automate the decomposition of complex model to simpler sub-models can be a very effective tool in performing the inference and uncertainty analysis of complex process models.

7.1.4 Investigating the "Hier-inf" approach

As discussed in previous chapters 4 and 5, the principle of "Hier-inf" approach was to isolate the data into series and perform the inference separately for each of the series. However, by doing so, it was difficult to ensure that the distributions of some of the fundamental parameters do not change when calibrated using different series of data. This problem, however, is quite general for any process model and not specific to the application studied, especially when the involved parameters are highly cross-correlated. One way to handle such a shortcoming of "Hier-inf" approach, is to be able to make the parameters independent by using some transformation techniques [60] in terms of the principal modes (variables) that can be determined using principle component analysis (PCA) [69]. In case of determining a solution to restrain the distribution of some parameters that are obtained from different series, the "Hier-inf" approach can be very effective in characterizing the

uncertainties involved in complex processes.

7.1.5 Propagation to techno-economic feasibility model

In this thesis, the propagation of uncertainties were performed to determine the impact on the prediction of process performance variables such as product quality etc. However, another key performance metric that would generally be of interest is the prediction uncertainty involved in the estimate of techno-economics. That would involve translating the dynamic simulation results to the time averaged values to be employed in the techno-economic analysis model [70]. Although, propagation by itself is not of a huge computational challenge, automating the translation of the dynamic process output to a cyclic average to be fed into the economic model is non-trivial task and can be pursued as one of the future research direction.

Appendices

APPENDIX A

ESTIMATION OF ψ FOR C803 SILICA

The values for the parameters along with their references are listed in Table.A.1.

Table A.1: Values of parameters for estimating ψ

Parameters	Symbols	Values
Fraction of total amine sites on surface	ψ^{sba-15} at 40 ^{o} C at 50 wt%	0.64 [55]
Surface area after loading (SBA-15)	A_l^{sba-15}	1000 m^2/g-sorbent [71]
Surface area before loading (SBA-15)	A_b^{sba-15}	80 m^2/g-sorbent [71]
Surface area before loading (C803)	A_b^{c803}	209 m^2/g-sorbent [11]
Surface area before loading (C803)	A_l^{c803}	37 m^2/g-sorbent [11]
Molecular weight of PEI segment	M_w	423 g/mol-PEI [71]
No of amine sites per mol of PEI	N	11
Total amine loading (C803)	M^{c803}	9.54 mmol N/g-sorbent [11]

The surface area lost on loading the silica with PEI is given by $A_b - A_l$, where A_b and A_l are the surface areas of the silica before and after loading the sorbent. Let S_p be the surface area occupied by a unit weight of surface PEI and w be the weight $\%$ sorbent loading in the silica. Now, the fraction of total amine sites (total PEI) available on the surface of the sorbent pores (surface PEI) can be determined as follows.

$$\psi = \frac{(A_b - A_l)}{S_p w} \tag{A.1}$$

Here S_p, the surface area occupied by a unit weight of surface PEI is the characteristic parameter for which the value is unknown. To determine the value of S_p, the data obtained

using temperature programmed desorption experiment by Wang et al [71] with the PEI loaded SBA-15 are used. S_p is given by the following equation and is calculated as 2875 m^2/g-surface PEI.

$$S_p = \frac{(A_b^{sba-15} - A_l^{sba-15})}{\psi^{sba-15} w^{sba-15}} \tag{A.2}$$

Knowing S_p, the fraction of total amine sites available on the surface of the sorbent pores can be calculated for any silica using Eq.A.1. Here, w for the silica particle of interest can be calculated using the following relation $w = MM_w/N$, as the measurements are available for the total amine loading in the sorbent M. As the amine loading is typically reported as mmol N (nitrogen)/g-sorbent, the conversion factors of molecular weight of PEI M_w and number of N(nitrogen) sites available per mol of PEI are used to obtain w. It is to be noted that the estimates are solely based on the value of ψ reported by Wang et al[55] for SBA-15 using temperature programmed desorption. Therefore, the value of ψ estimated may be subject to substantial uncertainty.

Appendices

APPENDIX B

DERIVATION OF OVERALL MASS TRANSFER COEFFICIENT

Mass balance for each of the different phases viz. gas, macropore and sorbent, is derived individually by assuming that the respective phase is the controlling resistance for each case. The flux across each of the phase is determined and equated with one another to find the overall resistance.

B.1 Sorbent

The mass balance for a single sorbent particle is given as follows.

$$(1-\epsilon_s)\frac{\partial q'}{\partial t}+\epsilon_s\frac{\partial C_{sg}}{\partial t}=\epsilon_s D_p\frac{1}{r}\frac{\partial^2}{\partial r^2}\Big(r^2\frac{\partial C_{sg}}{\partial r}\Big) \tag{B.1}$$

where q' is the CO_2 concentration in the solid phase within the sorbent and C_{sg} is the CO_2 cocentration within the micropores of the sorbent particle. The total concentration within the sorbent particle qs is defined as follows.

$$qs=(1-\epsilon_s)q'+\epsilon_s C_{sg} \tag{B.2}$$

Now, a quadratic profile is assumed for qs across the sorbent radius as given by

$$qs=A'+B'r^2, r\in(0,r_s) \tag{B.3}$$

where A' and B' are constants. The boundary conditions in the radial domain for the sorbent concentration are defined as follows.

$$qs = q_{ms}, r = r_s \tag{B.4}$$

$$\frac{\partial qs}{\partial r} = 0, r = 0 \tag{B.5}$$

Eq.B.1 is integrated over the sorbent volume to obtain the following equation

$$\frac{\partial q}{\partial t}\frac{4}{3}\pi r_s^3 = \epsilon_s D_p 4\pi\Big(r_s^2 \frac{\partial C_{sg}}{\partial r}|_{r=r_s}\Big) \tag{B.6}$$

where q is the volume average of qs. Similarly, the equation qs in Eq.B.3 is integrated over the volume as follows.

$$q = \frac{3}{4\pi r_s^3}\int_0^{r_s} 4\pi r^2 (A' + B' r^2) dr \tag{B.7}$$

$$= A' + (3/5)B' r_s^2 \tag{B.8}$$

Using both Eq.B.4 and Eq.B.8, we obtain B' as follows.

$$B' = (5/2)(q_{ms} - q)/r_s^2 \tag{B.9}$$

Appyling the chain rule,

$$\frac{\partial C_{sg}}{\partial r}|_{r=r_s} = \Big(\frac{\partial qs}{\partial r}/\frac{\partial qs}{\partial C_{sg}}\Big)|_{r=r_s} \tag{B.10}$$

Solving for numerator of right hand side in the above equation using Eq.B.3 and substituting for B' from Eq.B.9, we obtain as follows.

$$\frac{\partial C_{sg}}{\partial r}|_{r=r_s} = \frac{5(q_{ms} - q)}{r_s \partial q'/\partial C_{sg}(1 - \epsilon_s)} \tag{B.11}$$

In the above equation, the accumulation in the gas phase within the sorbent $\epsilon_s C_{sg}$ is neglected compared to the solid phase concentration q' while substituting for the denominator in Eq.B.10 using Eq.B.2. The above equation is back substituted in Eq.B.6 to get the following.

$$\frac{\partial q}{\partial t} = \frac{15\epsilon_s D_p}{r_s^2 \frac{\partial q'}{\partial C_{sg}}(1-\epsilon_s)}(q_{ms} - q) \tag{B.12}$$

In the above equation, the coefficient of driving force is the inverse of sorbent mass transfer resistance as given below:

$$K_s = \frac{15\epsilon_s D_p}{r_s^2 \frac{\partial q'}{\partial C_g}(1-\epsilon_s)} \tag{B.13}$$

B.2 Macropore

A similar approach to the sorbent mass transfer resistance is followed to obtain the macropore resistance. The total sorbent concentration in the fiber phase is defined as follows.

$$\zeta' = q v_s (1-\epsilon_s) + C_{mg}\epsilon_f \tag{B.14}$$

where ζ' is the total fiber phase concentration, C_{mg} is the gas phase concentration in the macropore and v_s is the sorbent loading in the fiber. A quadratic profile across the fiber annular cross section is assumed as follows.

$$\zeta' = A'' + B'' r^2, r \in (r_{ID}, r_{oD}) \tag{B.15}$$

The mass balance in the fiber phase is given by the following,

$$\frac{\partial \zeta'}{\partial t} = D_{fp}\epsilon_f \frac{1}{r}\frac{\partial}{\partial r}\Big(r\frac{\partial C_{mg}}{\partial r}\Big) \tag{B.16}$$

with the following boundary conditions,

$$\zeta' = \zeta_{ms}, r = r_{oD} \tag{B.17}$$

$$\frac{\partial \zeta'}{\partial r}|_{r=r_{ID}} = 0 \tag{B.18}$$

Similarly, Eq.B.16 is averaged across the fiber volume to get the volume averaged fiber phase concentration as following

$$\frac{\partial \zeta}{\partial t} = \frac{2D_{fp}\epsilon_f r_{oD}}{r_{oD}^2 - r_{ID}^2}\frac{\partial C_{mg}}{\partial r}|_{r=r_{oD}} \tag{B.19}$$

where ζ is the volume averaged concentration of ζ'. Eq.B.15 is averaged in volume to obtain the following.

$$\zeta = A'' + B''(r_{ID}^2 + r_{oD}^2)/2 \tag{B.20}$$

Applying the averaged over the volume of fiber to the Eq. B.14, we obtain the following:

$$\zeta = \bar{q}v_s(1-\epsilon_s) + \bar{C}_{mg}\epsilon_f \tag{B.21}$$

where, $\bar{q}$ and $\bar{C}_{mg}$ are the volume averaged sorbent concentration as $\bar{q} = 1/V_{fiber}\int_{r_{iD}}^{r_{oD}} 2\pi r L q dr$ and $\bar{C}_{mg} = 1/V_{fiber}\int_{r_{iD}}^{r_{oD}} 2\pi r L C_{mg} dr$ respectively.

Applying the chain rule, we have

$$\frac{\partial C_{mg}}{\partial r}|_{r=r_{oD}} = \Big(\frac{\partial \zeta'}{\partial r}/\frac{\partial \zeta'}{\partial C_{mg}}\Big)|_{r=r_{oD}} \tag{B.22}$$

Using both Eq.B.17 and Eq.B.15, we can solve for B'' as follows.

$$B'' = \frac{2(\zeta_{ms} - \zeta)}{(r_{oD}^2 - r_{ID}^2)} \tag{B.23}$$

Solving for the numerator in Eq.B.22 using Eq.B.15 and substituting for B'', we get the following.

$$\frac{\partial C_{mg}}{\partial r}|_{r=r_{oD}} = \frac{4r_{oD}(\zeta_{ms}-\zeta)}{(r_{oD}^2 - r_{ID}^2)\frac{\partial q'}{\partial C_{mg}}(1-\epsilon_f)vs(1-\epsilon_s)} \tag{B.24}$$

where again accumulation in the gas phase of macropores C_{mg} is neglected compared to q'. Back substitution of the above in Eq.B.19 yields the following equation.

$$\frac{\partial \zeta}{\partial t} = \frac{8D_{fp}\epsilon_f r_{oD}^2}{(r_{oD}^2 - r_{ID}^2)^2\frac{\partial q'}{\partial C_{mg}}(1-\epsilon_f)(1-\epsilon_s)vs}(\zeta_{ms}-\zeta) \tag{B.25}$$

In the above equation, the driving force $\zeta_{ms} - \zeta$ is modified as follows, using the averaged form of Eq.B.14.

$$\zeta_{ms} - \zeta = (q_{ms}v_s(1-\epsilon_f) - qv_s(1-\epsilon_f)) + (\epsilon_f C_{ms} - \epsilon_f \bar{C}_{mg}) \tag{B.26}$$

where $\bar{C}_{mg}$ is the average concentration in the fiber macropore gas phase. In the above, the terms involving the gas phase concentration C_{mg} and C_{ms} are ignored based on the assumption of negligible accumulation in macropore compared to the solid phase concentration q to obtain the following.

$$\zeta_{ms} - \zeta = (q_{ms} - q)v_s(1-\epsilon_f) \tag{B.27}$$

The above equation is rewritten in terms of gas phase concentration using the approximate equilibrium relation, $q = (\partial q/\partial C_{mg})\bar{C}_{mg}$, $q_{ms} = (\partial q/\partial C_{mg})C_{ms}$ as follows.

$$\zeta_{ms} - \zeta = \frac{\partial q}{\partial C_{mg}}(C_{ms} - \bar{C}_{mg})v_s(1-\epsilon_f) \tag{B.28}$$

Subsituting for ζ in Eq.B.25 using Eq.B.21 and neglecting the gas phase accumulation $\bar{C}_{mg}\epsilon_f$ in the sorbent pores, Eq.B.25 can be re-written using Eq.B.28 in the following man-

ner.

$$\frac{\partial q}{\partial t} = \frac{8D_{fp}\epsilon_f r_{oD}^2}{(r_{oD}^2 - r_{ID}^2)^2(1-\epsilon_f)v_s(1-\epsilon_s)}(C_{ms} - \bar{C}_{mg}) \tag{B.29}$$

In obtaining Eq. B.29, $\partial q/\partial C_{mg}$ is simplified using Eq.B.2 assuming $\partial q/\partial C_{mg} = \partial qs/\partial C_{mg}$. In Eq.B.29, the coefficient of the driving force is the inverse of the fiber phase mass transport resistance $K'_m = \frac{8D_{fp}\epsilon_f r_{oD}^2}{(r_{oD}^2 - r_{ID}^2)^2(1-\epsilon_f)v_s(1-\epsilon_s)}$.

B.3 Gas phase

Assuming gas phase as the controlling resistance in the fiber and negligible accumulation in the fiber macropores, the mass balance for fiber phase concentration can be written as following.

$$\frac{\partial q}{\partial t} = \frac{2k_g r_{oD}}{(r_{oD}^2 - r_{ID}^2)v_s(1-\epsilon_f)}(C_g - C_{ms}) \tag{B.30}$$

where the driving force is from the bulk gas phase concentration to the one at the outer radius of the fiber C_{ms}. In this equation, the coefficient of driving force is the inverse of the gas phase resistance $K'_g = \frac{2k_g r_{oD}}{(r_{oD}^2 - r_{ID}^2)vs(1-\epsilon_f)}$

B.4 Flux across interfaces

When all the resistances are controlling, the driving force gets modified accordingly in each of the phase equations such that the resistances act in series (Fig.5). The driving force across the gas film is $C_g - C_{ms}$ where C_{ms} is the gas phase concentration at $r = r_{OD}$. For fiber phase, the driving force is $C_{ms} - C_{mo}$ where C_{mo} is the gas phase concentration at the outer radius of sorbent in equilibrium with the sorbent concentraton q_{ms} at sorbent radius r_s. Finally, the mass transfer driving force across the sorbent phase is $q_{ms} - q$. Equating the flux across the interfaces to ensure no accumulation at the interface gives

$K'_g(C_g - C_{ms}) = K'_m(C_{ms} - C_{mo}) = K_s(q_{ms} - q) = K_{ov}(KC_g - q)$.

To do the conversion between the sorbent phase concentration and the gas phase concentration, the approximate equilibrium relation $q_{eq} = KC$ is used where $K = \frac{\partial q_{eq}}{\partial C}$. Performing the algebra gives the equation of the overall mass transfer coefficient in terms of the component coefficients.

$$K_{ov} = \frac{1}{K/K'_g + K/K'_m + 1/K_s} \tag{B.31}$$

Back substituting in to the mass balance, we get

$$\frac{\partial q}{\partial t} = \frac{1}{(\partial q_{eq}/\partial C_g)/K'_g + (\partial q_{eq}/\partial C_g)/K'_m + 1/K_s}(q_{eq} - q) \tag{B.32}$$

where q_{eq} is the sorbent concentration in equilibrium with the gas phase concentration C_g. We define, $K_g = K'_g/(\partial q_{eq}/\partial C_g)$ and similarly $K_m = K'_m/(\partial q_{eq}/\partial C_g)$ respectively. As the equilibrium capacity is reduced as ψq_{eq} and $(1-\psi)q_{eq}$ with respect to the sorbent phase resistances, we define $K_{s,k} = K_s/\psi$ and $K_{s,p} = K_s/(1-\psi)$ respectively.

REFERENCES

[1] T.J. Smith and L.A. Marshall. "Bayesian methods in hydrologic modeling: a study of recent advancements in Markov chain Monte Carlo techniques". In: *Water Resources Res.* 44 (2008), W00B05.

[2] W.N. Edeling et al. "Bayesian estimates of parameter variability in $k - \epsilon$ model". In: *Journal of Comput. Phys.* 258 (1 2014), pp. 73–94.

[3] A. Bouchard-Coete, S. Sankararaman, and M. Jordan. "Phylogenetic inference via sequential Monte Carlo". In: *Syst. Biol.* 61 (4 2012), pp. 579–593.

[4] R. Pouillot et al. "Estimation of uncertainty and variability in bacterial growth using Bayesian inference. Application to listeria monocytogenes." In: *Int J Food Microbiology* 81 (1 2003), pp. 87–104.

[5] S.H. Hsu et al. "Bayesian framework for building kinetic models of catalytic systems". In: *Ind. Eng. Chem. Res.* 48 (2009), pp. 4768–90.

[6] M.C. Coleman and D.E. Block. "Bayesian parameter estimation with informative priors for nonlinear systems". In: *AIChE J.* 52 (2006), pp. 651–67.

[7] J. Kalyanaraman et al. "Uncertainty quantification via Bayesian inference using sequential Monte Carlo methods for CO_2 adsorption process". In: *AIChE J.* 62 (9 2016), pp. 3352–68.

[8] P. Angelikopoulos, C. Papadimitriou, and P. Koumoutsakos. "Bayesian uncertainty quantification and propagation in molecular dynamics simulations: A high performance computing framework." In: *Journal of Chemical Physics.* 137 (2012), pp. 144103–10.

[9] D.C. Miller et al. "Carbon capture simulation initiative: a case study in multiscale modeling and new challenges". In: *Annual. Rev. Chem. Biomol. Chem. Eng.* 5 (2014), pp. 303–323.

[10] M.C. Kennedy and A. O'Haga. "Bayesian calibration of computer models". In: *J.R. Stat. Soc. B* 63 (2001), pp. 425–464.

[11] Y. Labreche et al. “Post-spinning infusion of poly(ethyleneimine) into polymer/silica hollow fiber sorbents for carbon dioxide capture”. In: *Chemical Engineering Journal* 221 (2013), pp. 166–175.

[12] Y. Fan et al. “Evaluation of CO_2 adsorption dynamics of polymer/silica supported poly(ethylenimine) hollow fiber sorbents in rapid temperature swing adsorption”. In: *International Journal of Greenhouse Gas Control* 21 (2014), pp. 61–71.

[13] G.D. Agostini. “Bayesian inference in processing experimental data principles and applications”. In: *Reports on Progress in Physics* 66 (2003), pp. 1383–1420.

[14] A. Gelman et al. *Bayesian data analysis, Third Edition*. John Wiley and Sons Inc., 2013.

[15] M. Alfaro, S. Zoller, and F. Lutzoni. “Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence.” In: *Molecular Biology and Evolution* ().

[16] K. Chaloner and I. Verdinelli. “Bayesian experimental design: a review.” In: *Statistical Science* 10 (3 1995), pp. 273–304.

[17] A.C. Atkinson, A.N. Donev, and R.D. Tobias. *Optimum experimental designs*. With SAS New York: Oxford University Press, 2007.

[18] G. Franceschini and S. Macchietto. “Model based design of experiments for parameter precision:state of the art”. In: *Chemical Engineering Sci.* 63 (2008), pp. 4846–72.

[19] R. Chu and J. Hahn. “Parameter set selection for estimation of nonlinear dynamic systems”. In: *AIChE J.* 53 (11 2008), pp. 2858–2870.

[20] H.N. Najm et al. “Uncertainty quantification in chemical systemss”. In: *International J. for Numerical Methods in Engg.* 80 (2009), pp. 789–814.

[21] Albrecht J. “Estimating reaction model parameter uncertainty with Markov chain Monte Carlo”. In: *Computers and Chemical Engg* 48 (2013), pp. 14–28.

[22] P. Angelikopoulos, C. Papadimitriou, and P. Koumoutsakos. “Bayesian uncertainty quantification and propagation in molecular dynamics simulations:A high performance computing framework.” In: *Journal of Chemical Physics.* 137 (2012), p. 144103.

[23] I. Anagu et al. “Estimating Freundlich isotherm parameters of heavy metals from multiple batch extraction tests using a Bayesian approach”. In: *Geoderma* 173 (2012), pp. 42–49.

[24] D.S. Mebane et al. “Bayesian calibration of thermodynamic models for the uptake of CO_2 in supported amine sorbents using ab initio priors”. In: *Physical Chemistry Chemical Physics* 15 (12 2013), pp. 4355–66.

[25] M.A. Duran and B.S. White. “Bayesian estimation applied to effective heat transfer coefficients in a packed bed”. In: *Chemical Engg. Science* 50 (1995), pp. 495–510.

[26] W.A. Lane et al. “Numerical modeling and uncertainty quantification of a bubbling fluidized bed with immersed horizontal tubes”. In: *Powder Technology* 253 (2014), pp. 733–743.

[27] D. Xiu and G.E. Karniadakis. “Modeling uncertainty in flow simulations via generalized polynomial chaos”. In: *Journal of Computational Physics* 187 (2003), pp. 137–167.

[28] Y. Marzouk and D. Xiu. “A stochastic collocation approach to Bayesian inference in inverse problems”. In: *Commun. Comput. Phys.* 6 (1 2009), pp. 826–847.

[29] A. Agarwal, L.T. Biegler, and S.E. Zitney. “Simulation and optimization of pressure swing adsorption systems using reduced-order modeling.” In: *Ind. Eng. Chem. Res.* 48 (2009), 23272343.

[30] J. Yuan et al. “Reliable multi-objective optimization of high speed WEDM process based on Gaussian process regression.” In: *Int J. Machine Tools and Manufacture* 48 (2008), pp. 47–60.

[31] K.J. Ryan. “Estimating expected information gains for experimental designs with application to random fatigue-limit model”. In: *Journal of Computatiomal and Graphical Statistics* 12 (3 2003), pp. 585–603.

[32] G. Terejanu, R.R. Upadhyay, and K Miki. “Bayesian experimental design for the active nitridation of graphite by atomic nitrogen.” In: *Experimental Thermal and Fluid Science* 36 (2012), pp. 178–193.

[33] S. Mosbach et al. “Iterative improvement of Bayesian parameter estimates for an engine model by means of experimental design.” In: *Combustion and Flame* 159 (2012), pp. 13303–13.

[34] P. Muller and G. Parmigiani. “Optimal design via curve fitting of Monte Carlo experiments”. In: *Journal of Americal Stat. Assoc.* 90 (1995), pp. 1322–1330.

[35] A. Solonen, H. Haario, and M. Laine. “Simulation-based optimal design using a response variance criterion.” In: *Journal of Computational and Graphical Statistics* 21 (1 2012), pp. 234–252.

[36] X. Huan and Y.M. Marzouk. “Simulation-based optimal Bayesian experimental design for nonlinear systems.” In: *Journal of Computational Physics* 232 (1 2013), pp. 288–317.

[37] P. E. Hadjidoukas et al. “Π4U: A high performance computing framework for Bayesian uncertainty quantification of complex models”. In: *Journal Comput. Physics* 284 (2015), pp. 1–21.

[38] M.D. McCay. “Latin hypercube sampling as a tool in uncertainty analysis of computer models”. In: *1992 Winder Sim.Conf.Proc, New York* (1992).

[39] V.R. Joseph et al. “Sequential exploration of complex surfaces using Minimum Energy Designs”. In: *Technometrics* 57 (‘ 2015), pp. 57–64.

[40] H Yang et al. “Progress in carbon dioxide separation and capture : A review”. In: *Journal of Environmental Sciences* 20 (2008), pp. 14–27.

[41] S. Choi, J.H. Drese, and C.W. Jones. “Adsorbent materials for carbon dioxide capture from large anthropogenic point sources”. In: *ChemSusChem* 2 (2009), pp. 796–854.

[42] R.P. Lively et al. “CO_2 sorption and desorption performance of thermally cycled hollow fiber sorbents”. In: *International Journal of Greenhouse Gas Controls* 10 (2012), pp. 285–294.

[43] J. Kalyanaraman et al. “Modeling and experimental validation of carbon dioxide sorption on hollow fibers loaded with silica-supported poly(ethylenimine)”. In: *Chemical Engg Journal* 259 (1 2014), pp. 737–751.

[44] F. Rezaei et al. “Modeling of rapid temperature swing adsorption using hollow fiber sorbents”. In: *Chemical Engineering Science* 113 (2014), pp. 62–76.

[45] J. Kalyanaraman et al. “Bayesian estimation of parametric uncertainties, quantification and reduction using optimal design of experiments for CO_2 adsorption on amine sorbents”. In: *Comput. Chem. Engg.* 81 (2015), pp. 376–388.

[46] Y. Labreche et al. “Direct dual layer spinning of aminosilica/Torlon hollow fiber sorbents with a lumen layer for CO_2 separation by rapid temperature swing adsorption”. In: *Journal of Applied Polymer Science* 132 (2015), pp. 41845–55.

[47] F. Rezaei et al. “Aminosilane-grafted polymer/silica hollow fiber adsorbents for CO_2 capture from flue gas”. In: *ACS applied materials and interfaces* 5 (2013), pp. 3921–31.

[48] P. Bollini, S.A. Didas, and C.W. Jones. "Amine-oxide hybrid materials for acid gas separations". In: *J. Mater. Chem.* 21 (2011), pp. 15100–20.

[49] Y. Fan et al. "Dynamic CO_2 adsorption performance of internally cooled silica-supported poly(ethylenimine) hollow fiber sorbents". In: *AIChE J.* 60 (11 2015), pp. 3878–87.

[50] J. Happel. "Viscous flow relative to arrays of cylinders". In: *AIChE Journal* 5 (1959), pp. 174–177.

[51] J. Welty et al. *Fundamentals of momentum, heat and mass transfer, 5th ed.* John Wiley and Sons, Inc, 2008.

[52] P. Bollini et al. "Dynamics of CO_2 adsorption on amine adsorbents. 2. insights into adsorbent design". In: *Ind. Eng. Chem. Res* 51 (2012), pp. 15153–62.

[53] D.M. D.M. Ruthven, S. Farooq, and K.S. Knaebel. *Pressure Swing Adsorption*. CRC Press, Chapman and Hall Inc, 1993.

[54] R. Serna-Guerrero and A. Sayari. "Modeling adsorption of CO_2 on amine-functionalized mesoporous silica. 2: Kinetics and breakthrough curves". In: *Chemical Engineering Journal* 161 (2010), pp. 182–190.

[55] X. Wang and C. Song. "Temperature-programmed desorption of CO_2 from polyethylenimine-loaded SBA-15 as molecular basket sorbents". In: *Catalysis Today* 194 (2012), pp. 44–52.

[56] A. Asimakopoulou and A. Karabelas. "CO_2 sorption and desorption performance of thermally cycled hollow fiber sorbents". In: *Journal of Membrane Science* 282 (2006), pp. 430–441.

[57] H. Fujita and A.J. Kishimoto. "Diffusion-Controlled Stress Relaxation in Polymers. II. Stress Relaxation in Swollen Polymers". In: *Journal of Polymer Science* 28 (1958), pp. 547–559.

[58] *gPROMS Model Developer Guide, v3.4.0 ed*. Process Systems Enterprise, Ltd., 2011.

[59] D.M. Bates and D.G. Watts. *Nonlinear Regression Analysis and Its Applications*. Wiley, 2007.

[60] J.J. Faraway. *Linear Models with R, Second Ed.* Chapman and Hall, 2014.

[61] Scipy.optimize.curvefit. *https://docs.scipy.org*.

[62] W. Hastings. “Monte Carlo sampling methods using Markov chains and their applications”. In: *Biometrika* 57 (1970), pp. 97–109.

[63] Y.F. Atchade and J.S. Rosenthal. “On adaptive Markov chain Monte Carlo algorithms”. In: *Bernoulli* 11 (2005), pp. 815–828.

[64] S. Wu et al. “Bayesian framework for force field selection in molecular dynamics simulations”. In: *Phil Trans R Soc A* 374 (2015), p. 20150032.

[65] P. Pernot and F. Cailliez. “A critical review of statistical calibration/prediction models handling data inconsistency and model inadequacy”. In: *arXiv:1611.04376* ().

[66] www.pace.gatech.edu. *PACE cluster*.

[67] V. Chen and M. Hlavacek. “Application of Voronoi tessellation for modeling randomly packed hollow-fiber bundles”. In: *AIChE Journal.* 40 (4 1994), pp. 606–612.

[68] H.P. Wynn. “Discussion of the paper by Kennedy and OHagan”. In: *Journal of the Royal Statistical Society, Series B* 63 (2001), pp. 450–451.

[69] B.M. Wise et al. “A theoretical basis for the use of principal component models for monitoring multivariate processes”. In: *Process control and quality* 1 (1990), pp. 41–51.

[70] S. Swernath et al. “Optimization and technoeconomic analysis of rapid temperature swing adsorption process for carbon capture from coal-fired power plant.” In: *Computer Aided Chem. Engg.* 36 (2015), pp. 253–278.

[71] X. Wang et al. “Molecular basket sorbents polyethylenimine-SBA-15 for CO_2 capture from flue gas: Characterization and sorption properties”. In: *Microporous and Mesoporous Materials* 169 (2013), pp. 103–111.

VITA

Jayashree Kalyanaraman, completed her undergraduate in chemical engineering in Alagappa College of Technology (ACTech), Anna University in 2005. After which, she pursued her Masters in Chemical Engineering in Indian Institute of Technology, Madras (IIT-M). During her masters program, she received the prestigious DAAD scholarship, which offered the opportunity to study and work for the master's thesis in Rheinisch-Westfaelische Technische Hochschule (RWTH), Aachen, Germany. Her masters thesis was on "Robust dynamic optimization under uncertainty for batch and semi-batch processes", which she worked under the advise of Prof. Wolfgang Marquardt. After graduating, she joined Dow Chemicals International Pvt Ltd in Pune, India and worked as a Process Research Engineer for over 3.5 years. In Spring 2012, she joined her PhD program in the School of Chemical and Biomolecular Engineering and pursued her doctoral research on "Performing uncertainty quantification of complex chemical processes with application to post-combustion carbon capture", advised by Prof. Matthew J. Realff and Prof. Yoshiaki Kawajiri. Her husband, Balaji Muralidharan, is also pursuing his doctoral research in School of Aerospace Engineering at Georgia Institute of Technology. She is a proud mother of a little one, named Rohan Balaji, who is little over 2 years old now.