M 2023

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

MODEL TOPOLOGY IDENTIFICATION AND GLOBAL PARAMETERS ESTIMATION FOR PURIFICATION OF BSA AND MYOGLOBIN

GUILHERME DA COSTA AMARAL
MASTER THESIS IN CHEMICAL ENGINEERING
PRESENTED TO THE FACULTY OF ENGINEERING
OF THE UNIVERSITY OF PORTO

Master in Chemical Engineering

Model Topology Identification and Global Parameters Estimation for Purification of BSA and Myoglobin

Master dissertation

of

Guilherme da Costa Amaral

Developed within the course of dissertation

held in

LSRE – LCM – Laboratory of Separation and Reaction Engineering – Laboratory of Catalysis and Materials, part of ALiCE- Associate Laboratory in Chemical Engineering

Supervisor at FEUP: **Prof. Alexandre Ferreira**

Co-supervisor at FEUP: **Prof. Ana Mafalda Ribeiro**

Co-supervisor at NTNU: **Prof. Idelfonso Nogueira**

July 2023

Acknowledgment

First and foremost, I would like to offer my deepest gratitude to all the people that made this dissertation a possibility.

To Professor Alexandre Ferreira, who gave me the opportunity to develop this dissertation under a team of excellent people. Thank you for the invitation, for all the help provided when work was not going as planned, and for the pleasant mood the Professor provided during troubling days. The knowledge and wisdom the Professor offered me were essential not only to the development of the work but to the well-being of this student. It was a pleasure working with you.

To Professor Ana Mafalda Ribeiro, whose presence would calm even the most agitated mind. Thank you for all the knowledge passed onto me, for introducing me to the people of lab E406, and for the guidance provided throughout the dissertation. Your kind words gave me the motivation I needed to keep pushing forward and achieve what I know I'm capable of. I am very grateful to have been your student.

To Professor Idelfonso Nogueira, who helped me immensely in the realization of this dissertation, even though we were hundreds of kilometres apart. Your kind words filled me with determination to keep improving and to keep pushing even if the results were not what I hoped for. Thank you for all the assistance and guidance throughout this journey, especially in the computational realm. Without the Professor, this work would not be complete.

I would also like to thank my colleagues at lab E406. Their assistance was indispensable to the development of this dissertation. Their availability to help me when I had doubts or when I didn't know what to do is something that I am extremely grateful for. Moreover, the friendly environment enabled me to be the best I can be. It was a pleasure working alongside such friendly and easy-going people.

I would also like to express my appreciation to those who did not assist me directly but provided support and consideration that allowed me to get through this long journey with my head held up high.

First, to my parents. Thank you for all the support you gave me all these years. You would always believe in me even when I doubted myself and offered me kind and helpful words when this feeble mind was at a loss. You were the light at the end of the tunnel and the bliss in the moments of elation. Without you, none of this would be possible. I am very proud of being your son.

To my sister, Maria, because even if we didn't see eye to eye, you were always there to make me laugh. Your presence would always clear my darkest days and your determination to fulfil

your dreams is something that I aspire to have. You will always be the pride of this older brother. Thank you for all the fun times we shared throughout our entire lives.

To my outer family, that supported me and my journeys to fulfil my dreams all these years. Your motivating words and care for my well-being were always something that I would appreciate from the bottom of my heart. Thank you for always being there.

To my closest friends, Gonçalo, Guilherme, and, Rui, whose presence was the beacon of joy I would seek. Even though you gave me headaches all the time, your friendship is something that I will value for the rest of my life. The moments and the conversations we shared will always be the best memories one could ever have. Thank you for always being beside me on this difficult journey.

To the rest of the friends that I gained throughout this 5-year path, thank you for all the kind words and for the fun times we shared. Your company was something that I was looking forward every day. Special thanks to my friend Diana, who accompanied me in this dissertation journey, through every difficult time and through times of joy. Thank you for the company, both in lab E406 and during the lunch breaks.

To my roommates throughout the years and my best friends, Diogo and João, which gratitude I cannot express in words, but nevertheless try. You took me in as a friend when no one else would, and shaped the person I am today. The times we were together, the conversations we had, and the experiences we shared were the best moments one could ever wish for. From the bottom of my heart, thank you. Our friendship is something I won't forget and will be the fuel of my motivation and determination for many years to come.

And finally, to the friends I left behind, in my hometown, even though we don't see each other as much, our late-night calls were something I was always looking for. They were filled with useless topics but cleansed any unpleasant thoughts. Thank you for all the consideration and concern during hard times and for your presence during the best ones.

I extend my gratitude to the Norwegian University of Science and Technology, Department of Chemical Engineering, Process Systems Engineering group for all the help provided throughout the dissertation.

Alexandre Ferreira, supervisor of this dissertation, is a member of LSRE-LCM - Laboratory of Separation and Reaction Engineering – Laboratory of Catalysis and Materials, financially supported by LA/P/0045/2020 (ALiCE), UIDB/50020/2020 and UIDP/50020/2020 (LSRE-LCM), funded by national funds through FCT/MCTES (PIDDAC).

Abstract

Proteins are biological compounds that are vital in both biological systems and industrial applications. Several protein separation techniques have been developed to better study and apply these nutrients. Among these techniques, chromatography separation, employed explicitly in a simulated moving bed (SMB), takes emphasis in this study. It is important to understand the mechanisms of protein separation; therefore, experimental work to separate the model proteins, myoglobin (Mb) and bovine serum albumin (BSA), using three different materials, was performed prior to the development of the current dissertation. The current dissertation aimed to utilize this study to estimate the mathematical model parameters that better align with the experimental data obtained before, for a specific case study, where the separation was performed using Sephadex G-50.

To achieve that goal, adequate objective functions were formulated to be employed in Particle Swarm Optimization (PSO) and, therefore, derive a set of parameters that would minimize the residual between the mathematical models and the experimental data. Initially, the optimization was applied exclusively for the fixed bed experiments to obtain reference parameters that would validate further estimations. Subsequently, an estimability analysis was performed to quantify the correlation between the parameters and their impact on the generated mathematical model. The total parameter estimation, encompassing all experiments available, including SMB experiments, was then performed. Finally, a simple uncertainty evaluation was carried out.

After the acquisition of reference parameters, the estimability analysis results showed that the mass transfer coefficients were parameters that should not be estimated and were excluded from further estimations. Parameter search bounds, defined for the PSO, were adjusted for the total parameter estimation after a particle density region analysis. The total parameter estimation yielded a set of parameters that created a mathematical model with a satisfactory fit to the SMB experimental data and Mb fixed bed experimental data while revealing an exceptional adjustment to the BSA and binary fixed bed experiments. Comparing this set of parameters with the set of reference parameters, it was possible to observe that the main deviation occurred in the axial dispersion parameter, in the Mb and BSA fixed-bed experiments, and the BSA size exclusion constant parameter, in the binary fixed-bed experiment. Nevertheless, the total parameters obtained were deemed satisfactory for the purpose of this work. Lastly, the uncertainty evaluation was carried out with success.

Keywords: Proteins, Adsorption, Parameter Estimation, Particle Swarm Optimization, Estimability Analysis

Resumo

As proteínas são compostos biológicos importantes tanto para os sistemas biológicos como para as aplicações industriais. Várias técnicas de separação de proteínas têm sido desenvolvidas para melhor estudar e aplicar estes componentes. Entre estas técnicas, a separação por cromatografia, utilizada explicitamente num Leito Móvel Simulado (SMB), tem um destaque importante nesta dissertação. É importante compreender os mecanismos de separação das proteínas, pelo que foram realizados trabalhos experimentais de separação das proteínas modelo, mioglobina (Mb) e albumina de soro bovino (BSA), utilizando três materiais diferentes, antes do desenvolvimento da presente dissertação. A presente dissertação teve como objetivo utilizar esse estudo, estimando os parâmetros do modelo matemático que melhor se alinham com os dados experimentais obtidos anteriormente, para um caso de estudo específico, em que a separação foi realizada utilizando o gel *Sephadex G-50*.

Para atingir esse objetivo, foram formuladas funções objetivo adequadas para serem utilizadas em *Particle Swarm Optimization* (PSO) e, assim, derivar um conjunto de parâmetros que minimizassem a diferença entre os modelos matemáticos e os dados experimentais. Inicialmente, a otimização foi aplicada exclusivamente às experiências em leito fixo para obter parâmetros de referência que validassem estimações futuras. Posteriormente, foi efetuada uma análise de estimabilidade para quantificar a correlação entre os parâmetros e o seu impacto no modelo matemático gerado. A estimação total dos parâmetros, englobando todas as experiências disponíveis, incluindo as experiências SMB, foi então efetuada. Finalmente, foi desenvolvida uma avaliação simples da incerteza dos parâmetros.

Após a aquisição dos parâmetros de referência, os resultados da análise de estimabilidade mostraram que os coeficientes de transferência de massa eram parâmetros que não se apresentavam estimáveis e foram excluídos de futuras estimações. Os limites de pesquisa de parâmetros, definidos para o PSO, foram ajustados para a estimação dos parâmetros totais após uma análise das regiões de densidade de partículas. A estimação total produziu um conjunto de parâmetros que criou um modelo matemático com um ajuste satisfatório aos dados experimentais de SMB e aos dados experimentais de Mb em leito fixo, e revelou um ajuste excecional às experiências de leito fixo de BSA e de leito fixo para mistura binária. Comparando este conjunto de parâmetros com o conjunto de parâmetros de referência, foi possível observar que o principal desvio ocorreu no parâmetro de dispersão axial, nas experiências de Mb e de BSA em leito fixo, e no parâmetro constante de exclusão de tamanho para a proteína de BSA, na experiência de leito fixo para mistura binária. No entanto, os parâmetros totais obtidos foram considerados satisfatórios para o objetivo deste trabalho. Por último, a análise de incerteza foi realizada com sucesso.

Declaration

I, Guilherme da Costa Amaral, hereby declare, under word of honour, that this work is original and that all non-original contributions is indicated and due reference is given to the author and source

Guilherme da Costa Amaral

Porto, 3rd of July of 2023

Index

List of Figures

List of Tables

Notation and Glossary

A	Area of the fixed-bed column's cross-section	$\boldsymbol{m^2}$
C_f	Learning factors in Particle Swarm Optimization	-
C_i	Concentration in the fluid phase of component i	$\boldsymbol{g \cdot l^{-1}}$
C_i^F	Concentration of component i at the feed inlet of the fixed-bed column	$\boldsymbol{g \cdot l^{-1}}$
D_{ax}	Axial Dispersion Coefficient	$\boldsymbol{m^2 \cdot s^{-1}}$
G^{best}	Global best position for the particles in Particle Swarm Optimization	-
H	Hessian Matrix	-
K	Coverage Factor	-
k_h	Mass Transfer Coefficient	$\boldsymbol{s^{-1}}$
K_{SEC}	Size Exclusion Constant	-
L	Column Length	$\boldsymbol{m}$
$M(p)$	Magnitude of column p of the sensitivity matrix	-
N	Population size	-
N_{DV}	Number of Decision Variables	-
N_{iter}	Number of Iterations	-
N_p	Number of Particles	-
N_t	Number of chosen time-steps	-
N_y	Number of output variables	-
P^{best}	Particle best position in Particle Swarm Optimization	-
Q	Flowrate in an SMB equipment	$\boldsymbol{l \cdot min^{-1}}$
Q_{feed}	Feed Flowrate in the fixed bed column	$\boldsymbol{m^3 \cdot s^{-1}}$
$\bar{q}_i$	Average adsorbed concentration of component i	$\boldsymbol{g \cdot l_{solid}^{-1}}$
q_i^*	Adsorbed concentration of component i at equilibrium	$\boldsymbol{g \cdot l_{solid}^{-1}}$
R	Residual matrix	-
r^*	Random number between 0 and 1	-
S	Sensitivity Matrix	-
S_f	Sensitivity Matrix with only estimable parameters	-
S_i	Parameter i of the Sensitivity matrix	-
S_i'	Sensitivity vector for parameter i	-
S_{max}	Column of the sensitivity matrix with the highest magnitude	-
S_o	Orthogonalization of the sensitivity matrix in relation to S_{max}	-
S_p	Column p of the sensitivity matrix	-
t	Time	$\boldsymbol{s}$
t_n	Time instant evaluated	-
U_c	Propagated uncertainty	-
u_i	Interstitial Velocity	$\boldsymbol{m \cdot s^{-1}}$
u_p	Uncertainty of parameter p	-
$v_{i,k,n}$	Velocity of particle k at iteration i for decision variable n	-
v_{max}	Maximum velocity of the particles in Particle Swarm Optimization	-
x	Vector with state variables	-
X_i	Vector with the best parameters obtained by the Particle Swarm Optimization	-
$x_{i,k,n}$	Position of particle k at iteration i for decision variable n	-
$x_{i,k,n}^*$	Proposed position of particle k at iteration i for decision variable n	-
x_{max}	Upper bound of the particles in Particle Swarm Optimization	-
x_{min}	Lower bound of the particles in Particle Swarm Optimization	-
y	Output vector	-
$\bar{y}$	Mean of the parameter values obtained during Particle Swarm Optimization	-

y_{exp}	Matrix with experimental data	-
y_i	i-th output obtained	-
y^{PSO}	Values yielded by the Particle Swarm Optimization for a certain parameter	-
y_{sim}	Matrix with simulated data	-
z	Axial Direction	$\boldsymbol{m}$

Greek Letters

ε_b	Bed Porosity
θ	Parameter Vector
ω_0	Inertia Weight in Particle Swarm Optimization
ω_0^*	Updated value for the Inertia Weight in Particle Swarm Optimization

Indexes

$I, \dots, IV$	Section of the SMB equipment
D	Desorbent (Eluent) Stream (SMB)
E	Extract Stream (SMB)
F	Feed Stream (SMB)
i	Component
i	Iteration (PSO)
in	Node entrance
k	Column Number
k	Particle Number (PSO)
n	Decision Variable (PSO)
out	Node Exit
R	Raffinate Stream (SMB)

List of Acronyms

BSA	Bovine Serum Albumin
CPU	Central Processing Unit
HPLC	High Pressure Liquid Chromatography
IUPAC	International Union of Pure and Applied Chemistry
LDF	Linear Driving Force
Mb	Myoglobin
MF	Microfiltration
NF	Nanofiltration
PSO	Particle Swarm Optimization
RAM	Random-Access Memory
SMB	Simulated Moving Bed
TMB	True Moving Bed
UF	Ultrafiltration

1 Introduction

1.1 Framing and presentation of the work

Proteins are crucial in facilitating various biological functions within the human body. In addition, they are recognized to be one of the most important nutrients due to their ability in building cells and new tissue, promoting the repair of the biological system. Moreover, proteins are widely used in industrial applications such as protein-based nanocarriers, film coaters, composites, hydrogels, microparticles, beds, and even pest control[1]. To gain a comprehensive understanding of their effects, protein separation, isolation, and purification techniques have extreme importance and have been subject to extensive investigations.

The choice of techniques used for protein separation depends on the proteins targeted for separation. Various commonly used techniques include chromatography, dialysis, ultrafiltration, general electrophoresis, among others[2]. A thorough understanding of the mechanisms behind the process must be acquired to perform the separation and achieve an optimal setup of the equipment. In the context of chromatography separation, particularly in a simulated moving bed (SMB), the fundamental understanding of the adsorption mechanisms is crucial. It can explain how the proteins interact with the matrices and materials employed in the SMB columns. Numerous mathematical models can describe these adsorption mechanisms[3], and their characteristic parameters can be easily determined by experimental work. However, uncertainties and errors present in such experiments may lead to slightly inaccurate parameter values.

To aid the understanding of the underlying adsorption mechanisms and to refine the mathematical models that describe them, optimization and parameter estimation can be employed alongside this experimental work. These approaches allow for a more precise alignment of the models to the experimental data, enabling better prediction and enhanced understanding.

The basis of this dissertation is a Ph.D. thesis where the experimental protein separation was performed[4]. The present study aims to evaluate the assumed mathematical adsorption models employed in separating these proteins and to estimate a set of parameters that create a mathematical model with superior adjustment to the experimental data, by developing adequate objective functions for the problem at hand. For the estimation, Particle Swarm Optimization (PSO) will be employed. Additionally, further analysis, such as estimability and uncertainty analysis, will also be carried out in order to assist in enhancing and validating the obtained results.

1.2 Contribution of the author to the work

The protein separation and the development of the gPROMS code for the adsorption mechanism's mathematical models were originally developed by Albertina Rios as part of her Ph.D. thesis[4]. The equipment setup used for this separation and the gPROMS code that described the operation of this equipment were developed by Sá Gomes *et al.*[5] For this dissertation, I modified the adsorption and global SMB operation models in gPROMS to suit a parameter estimation study.

My primary contribution to this dissertation was the development of the MATLAB® code functions that enabled parameter estimation. Additionally, I built and formulated the objective functions that optimized this estimation process. The employed Particle Swarm Optimization (PSO) code was developed and provided by Idelfonso Nogueira; however, certain modifications were necessary in order to enable parameter estimation. I carried out these adaptations of the code myself. Moreover, the estimability analysis and uncertainty evaluation were also carried out by me.

1.3 Organization of the dissertation

This dissertation is divided into six main chapters.

Chapter 1 provides an overview of the work to be performed, including information regarding the author's contributions and an outline of the dissertation's structure.

Chapter 2 presents a context background and a state-of-the-art overview, offering a fundamental understanding of the subject at hand. Section 2.1 provides a basic description of proteins, while section 2.2 explores some separation and purification strategies employed for these compounds, ultimately culminating in chromatography, discussed in section 2.3. Section 2.4 deals with some essential knowledge about mathematical adsorption models, while Section 2.5 has a small introduction to the relevant equipment, specifically the simulated moving bed. Section 2.6 provides a vital understanding of several analyses employed in the parameter estimation. Finally, section 2.7 offers a small literature overview that includes similar works.

Chapter 3, focuses on the materials used in this work and the methodology employed. To better understand the methodology, this chapter covers the models used, the particle swarm optimization algorithm, the formulation and structure of the objective functions, the estimability analysis and the uncertainty evaluation.

Chapter 4 displays the various results obtained across the studies performed, highlighting the outcomes of the various steps, while Chapter 5 summarizes the conclusions drawn from these studies. Chapter 6 offers a brief assessment of the work done, emphasizing future work.

2 Context and State of the art

2.1 Proteins

Proteins are found in every living organism and play a crucial role in the biochemical reactions occurring within the biological systems. Proteins comprise 20 different amino acids and exhibit a wide range of structural diversity and complexity. They may contain a small group of each amino acid, while others may lack one or more of these 20 organic compounds. Therefore, variations in size and complexity are widespread among proteins[6]. Their function heavily depends on both of these factors since the three-dimensional structure of the macromolecule and the sequence of amino acids collectively determine the characteristic properties of each protein.

Generally, proteins can be considered to have four levels of structure: primary, secondary, tertiary, and quaternary. The primary structure refers to the linear arrangement of amino acids. Short linear chains are known as peptides, while longer chains are usually referred to as polypeptides. Polypeptides that possess well-defined three-dimensional structures are commonly referred to as proteins. The secondary structure represents the spatial arrangement resulting from the folding of some segments of the principal sequence. When hydrogen bonds form between certain parts of the amino acid chain, the protein can be shaped into distinct secondary structures such as α helix, β sheet, or a short U-shaped turn. A single protein may show various types of secondary structures. The tertiary structure refers to the overall conformation/three-dimensional structure of the residual chains. The tertiary structure is mainly stabilized by the non-polar side chains that exert hydrophobic bonds, by the polar residual chains, through hydrogen bonds, and the peptide bonds. Since these interactions are weak, the tertiary structure of a protein is not entirely rigid and usually undergoes small fluctuations. Lastly, quaternary structure refers to the number and relative positions of the proteins in an assembly of multiple polypeptide chains that form multimeric structures. Proteins can also associate with other proteins to achieve the highest level of protein structure, forming macromolecular assemblies and further diversifying their function roles[7].

Proteins exhibit numerous roles within the human body, where they contribute to cellular growth and tissue development, and in industrial applications. Hence, protein separation and purification processes are extremely important to aid and facilitate the study and application of these essential nutrients. Some examples of industrial applications include plant-based protein applications ranging from supplements in food products and edible coats/films to stabilizers/emulsifiers, adhesives, and hydrogels[8]. Proteins can also find significant applications

in the pharmaceutical industry as research, diagnostics, and vital therapeutic tools[9]. Membrane proteins, which are present in the human body and facilitate communication and transport across cell membranes, have received considerable attention due to their selectivity and specificity, making them valuable assets in membrane processes. Membrane-based industrial applications, such as biosensors, energy harvest, and water purification, have been the highlight of recent studies[10].

2.2 Protein Separation and Purification Strategies

Protein separation usually encompasses various stages with distinct objectives. These stages can vary in number, but their theme remains consistent, typically including recovery, concentration, purification, and product formulation. In the recovery stage, the primary objective is to separate the protein from the remaining components in a mixture by physical means such as filtration and centrifugation or by direct extraction. The concentration step aims to increase the protein concentration to enable further manipulation. The outcome of the purification step heavily relies on the technique used. This part of the separation methodology usually consists of three main sub-steps: capture, intermediate purification, and polishing. Capture involves isolating and concentrating the target proteins, while intermediate purification focuses on removing bulk impurities and main contaminants present in the mixture. Lastly, the polishing step is usually performed to eliminate some leftover impurities, protein aggregates, or residual contaminants. Finally, the product formulation stage is the step where the protein is transformed into the desired product for the final costumer[11].

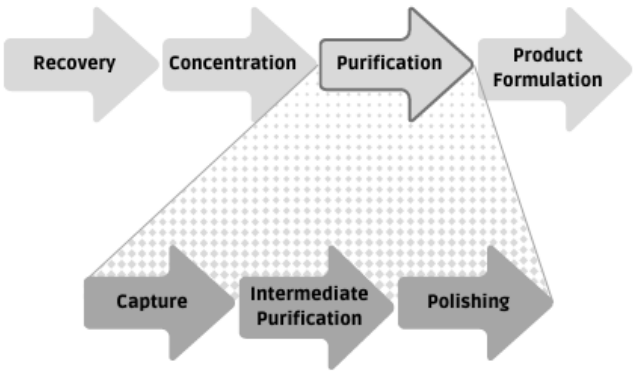

Figure 1 - Separation steps and sub-steps in a purification strategy

To purify the proteins, several types of methodologies may be employed. Among these methodologies, membrane techniques revealed satisfactory results. Microfiltration (MF), ultrafiltration (UF), and nanofiltration (NF) have shown promising results in protein separation and purification[12]. Chromatography has been shown as a superb purification strategy due to its

versatility and adaptability to the different properties of the proteins, and, therefore, is widely used today[13]. Other separation methods include dialysis and electrophoresis[2]; however, the former is rarely used on its own since it rarely achieves the desired purities. It is often combined with other purification methods so that it may yield the desired results[14].

Numerous studies have focused on improving protein separation methods to meet industrial requirements. Since membrane processes, despite their advantages, often lack high selectivity for particles of similar size, recent research has been released showcasing new technologies and methods, including operating at countercurrent conditions in hollow fiber membranes and the utilization of different polymers and additives for the membranes, in order to overcome such limitations[15-17].

For protein purification, recommended chromatography techniques include size-exclusion chromatography, affinity chromatography, ion-exchange chromatography, and liquid chromatography, such as high-pressure liquid chromatography (HPLC). It is important to note that achieving the desired purification level necessitates careful adjustment of parameters such as protein concentration in the initial mixture, operating conditions, flow rate, column dimensions, solid matrix, and fluid phase[18, 19].

2.3 Chromatography

The concept of chromatography, which involves the separation of components in a column, was introduced by Russian botanist Mikhail Tswett in 1903, and the term itself was coined in 1906[20]. Indeed, in chromatography, the sample is introduced into a column or a carrier containing a stationary phase (solid phase). The mobile phase transports the sample along the column. Due to the varying affinities of the sample components to the stationary phase, they travel through the vessel at different velocities, resulting in different retention times – the time between the sample introduction and the elution of each species from the column. The difference in retention times enables the separation of the components to occur[21].

There are several types of chromatography distinguished by the stationary phase, the mobile phase, and the components to be separated. The stationary phase is typically composed of a solid phase or a layer of a liquid adsorbed onto a solid support, while the mobile phase is either a liquid or a gas. A more general characterization indicates three different types of chromatography: liquid, gas, and super-critical chromatography. Numerous specific types of chromatography available include column chromatography, ion-exchange chromatography, gel permeation chromatography, affinity chromatography, paper chromatography, thin-layer chromatography, dye-ligand chromatography, hydrophobic interaction chromatography,

pseudo-affinity chromatography, high-pressure liquid chromatography (HPLC), and hydrophilic chromatography[22, 23].

2.4 Adsorption mathematical models

The development of mathematical models that describe the adsorption of components in an adsorbent is necessary to better understand the phenomena that occur in any chromatography. There are four key components required to model liquid-solid adsorption: bulk phase mass transport/transfer (convective and molecular diffusion), diffusion across the film to the adsorbent surface, intraparticle mass transfer (pore diffusion and surface diffusion), and adsorption/desorption equilibrium[24]. The adsorption equilibrium is often described by isotherms that can be distinct from one another due to the different solid phases used and the materials to be separated.

By solving a set of defined partial differential equations, coupled with algebraic ones, in which isotherms are included, it is possible to predict, for example, the breakthrough curve of a species and gain insight into the adsorption data, such as the adsorption mechanisms, the maximum adsorption capacity as well as the properties of the used adsorbents. Empirical isotherms models, such as the Linear, Freundlich, Redlich-Peterson, Sips, Toth, and Temkin isotherms, provide useful representations of the adsorption mechanisms. Additionally, chemical adsorption models like the Langmuir and Volmer isotherm models and physical adsorption models like the BET and Aranovich, also offer extensive understanding of the adsorption process. Ionic Exchange models are also a possibility in adsorption[25]. Furthermore, according to the International Union of Pure and Applied Chemistry (IUPAC), all these models can be separated into six types of isotherms, depending on their adsorption type[3]:

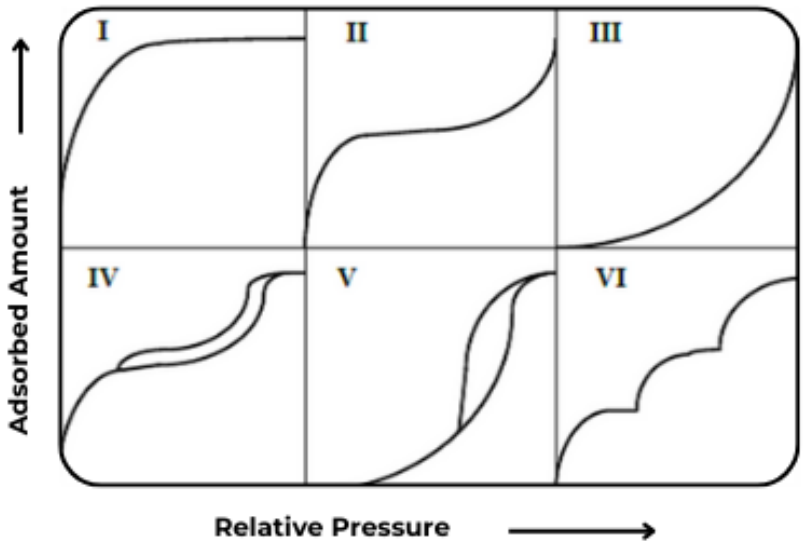

Figure 2 - Isotherm types and their respective graphical representation

Type I – Given by microporous solids with small external surfaces. The accessible micropore volume gives the limiting uptake;

Type II – Isotherm obtained from a non-porous or microporous adsorbent where the adsorption is mono or multilayer and unrestricted.

Type III – Gradual curvature is observed. These isotherms are uncommon.

Type IV – Isotherm characterized by its hysteresis loop, associated with condensation occurring in mesopores. The initial part of the isotherm is yielded by the mono/multi-layered adsorption. Many industrial mesoporous adsorbents follow this type of isotherms.

Type V – Very uncommon, and related to Type III. Obtained from specific porous adsorbents.

Type VI – Isotherms where various steps can be seen. The magnitude of the steps depends on the system and temperature and represents multilayer adsorption with various steps on a uniform non-porous surface. The step height represents the monolayer capacity for each adsorbed layer.

In liquid chromatography, several isotherm types are commonly encountered, including Linear (Henry), Langmuir, Freundlich, Temkin, and Sips isotherms. Besides the linear model, Langmuir is the most common. This model is based on the fact that the maximum adsorption corresponds to the saturation of the monolayer of the adsorbate molecules on the adsorbent surface, resulting in an isotherm with a characteristic shape resembling Type I isotherms[26]. In cases where multiple adsorption sites with varying energy distributions are present, a Bi-Langmuir isotherm can better describe the adsorption mechanism[27]. In highly diluted solutions, the adsorption equilibrium typically follows a linear model.

2.5 Simulated Moving Bed (SMB)

The first mention of the Simulated Moving Bed (SMB) technique came from a 1963 United States patent[28], used primarily for the separation of *p*-xylene from alkyl-aromatic C_8, and for the separation of paraffin in the petrochemical industry. Over time, SMB technology has downscaled and has been applied to the separation of smaller particles, including chiral molecules and finer chemicals[29].

SMB is an equipment where the two phases, liquid and solid, have a continuous, countercurrent movement facilitating the separation of the adsorbates in a sample. The operation of a simulated moving bed can better be understood by using its equivalent counterpart, the True Moving Bed (TMB). The term “True” signifies the presence of actual solid phase circulation. TMB enables the continuous chromatographic separation of components in a binary mixture, resulting in two outlet streams containing pure components: the extract and the raffinate. The

feed inlet and desorbent/eluent inlet contribute to the formation of these two streams[30]. The more retained compound exits the system via the extract stream, while the less retained compound exits through the raffinate stream. The affinities of the components to the solid phase, determined by their properties, dictate their retention. The four streams mentioned before divide the TMB into four sections (I to IV). Zone I lies between the desorbent inlet and the extract outlet, Zone II is located between the extract outlet and the feed inlet, Zone III exists between the feed inlet and the raffinate outlet, and, lastly, Zone IV extends from the raffinate outlet to the solid recycling. Figure 3 illustrates this zone division. The more retained component is adsorbed in Zone III, transported with the solid to Zone I for desorption, and then carried into the extract stream by the fluid phase. Simultaneously, the solid is regenerated and recycled in Zone IV. The less retained component can be adsorbed in Zone IV to allow for eluent regeneration, desorbed in Zone II to prevent extract stream contamination, and subsequently transported by the mobile phase to Zone III and lastly to the raffinate outlet stream[31].

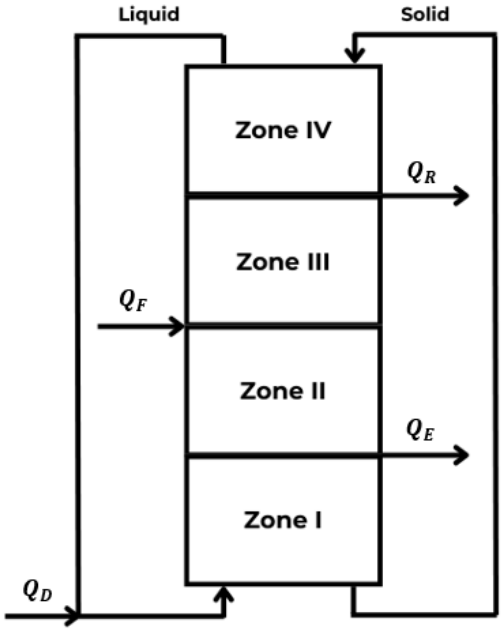

Figure 3 - Diagram of a TMB unit

These operation conditions yield favourable results; however, the flow of the solid phase in the TMB can lead to challenges such as mechanical erosion of the adsorbent, equipment abrasion, and even difficulties in maintaining the solid phase plug flow. Multiple columns are arranged in series to overcome these issues, where the solid phase remains stationary. The movement of the solid can be simulated by performing a shift of the inlet and outlet streams in the direction of the fluid flow within a predetermined time. Consequently, the solid phase does not physically move but simulates a countercurrent flow relative to the fluid phase, giving rise to the name "Simulated Moving Bed" (SMB). The results obtained in an SMB are comparable to those achieved

by operating a TMB under similar conditions[32]. Figure 4 illustrates an example of an SMB equipment:

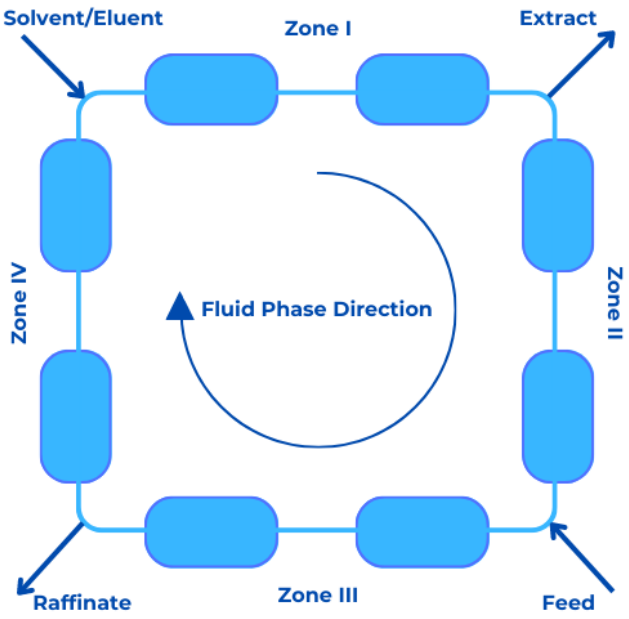

Figure 4 - Diagram of an SMB unit

Even though Figure 4 presents an SMB with two columns per section/zone, this is not factual in every equipment unit. An SMB process can have different operating modes depending on the dynamic configuration, flow rate modulation, concentration modulation, gradient operation, and the specific SMB configuration[33].

2.6 Parameter Estimation

2.6.1 Parameter Estimation Methods

To estimate the parameters of the mathematical models that characterize the SMB process, various methodologies may be employed. Analytical methods are feasible; however, they often prove to be very complex and impractical in engineering. Consequently, optimization problems can then be solved by deterministic or heuristic methods.

Deterministic methods guarantee convergence into an optimal solution, contrary to heuristic methods. However, their effectiveness heavily depends on the initial vector given, which means a precise initial guess will lead to more accurate results. In contrast, an imprecise initial vector can yield an unsatisfactory solution stuck in a local optimum, hindering problem resolution[34]. These approaches are then especially useful when dealing with simpler engineering optimization problems due to their quick convergence and result accuracy. However, in complex optimization problems, such as nonconvex or large-scale problems, deterministic

techniques may require a long time to converge or fail to converge. In such cases, heuristic approaches are recommended due to their flexibility[35].

The use of heuristic methods is very promising when estimating parameters in complex engineering processes, as they do not rely on initial guesses or conditions[34]. While heuristic methods do not necessarily reach the optimal solution, they strive to identify a solution as close as possible to this optimum. This significantly reduces the time required to find a solution to complex engineering problems and enables solving problems where the application of exact/analytical or deterministic methods is not possible[36]. Although solving time is reduced when using heuristic methods, the computing power needed is much higher when compared to other approaches. This may not pose a problem in simpler optimization problems. However, as the complexity of the problem increases, the required computing power also escalates[37]. There are a plethora of different heuristic techniques and the majority of them leverage computational power to estimate and optimize parameters in various processes. Examples of recently utilized and studied methods include swarm intelligence and evolutionary computation[38]. Swarm intelligence will be the chosen tool employed in this work, especially Particle Swarm Optimization.

2.6.2 Particle Swarm Optimization (PSO)

Particle Swarm Optimization was pioneered by Kennedy and Eberhart in 1995, and it is described as a technique for continuous nonlinear function optimization, drawing inspiration from swarm theory, evolutionary computation, and genetic algorithms[39]. A Particle Swarm Optimization method manipulates a defined set of particles to search for the optimal solution using not only the individual memory of each particle but also the collective knowledge of the entire swarm.

Initially, these particles are randomly distributed within the design space. Velocity vectors are then calculated for each particle, considering its original position. With this information, the position of each particle can be updated accordingly. These steps are repeated, which enable the swarm of particles to explore promising regions that offer optimal solutions to the given problem[40].

Several additions and modifications have been proposed to improve the efficiency of the basic PSO algorithm. The first method developed by Kennedy and Eberhart was not very efficient, so in the algorithm's infancy, the parameter inertia weight, ω_0, was introduced. This parameter plays a crucial role in balancing the local and global searches, preventing the particles from converging into a local optimum when the local search is predominant and preventing the excessive movement of the particles when the global search is dominant. It should be noted that lower inertia weight favours the local search, while higher inertia weight favours the global

search[41]. Moreover, the particles can be programmed to only store feasible solutions in their memory, simplifying the algorithm and yielding satisfactory results in less time[42]. Other improvements to the PSO algorithm involve refining the position and velocity update strategies, enhancing the particle learning process, modifying constrain methods, combining PSO with other techniques, and introducing mechanisms to maintain population diversity[43]. By incorporating these enhancements, PSO can be adapted to various optimization problems, providing more efficient and effective solutions.

2.6.3 Estimability analysis

Estimability analysis is a valuable technique employed to explore correlations between parameters, quantify their influence, and simplify any given optimization problem. Estimability methods assess the viability of estimating certain parameters based on the system input and output data. This step is a crucial stage in parameter estimation, particularly in complex problems, as parameter correlation often occurs and can inhibit successful estimation[44, 45].

Estimability analysis was pioneered by Yao *et al.* using a methodology based on orthogonalization. Their approach reported adequate results when addressing the field of parameter estimation for complex systems with a large number of parameters[46]. This methodology effectively reduced the number of parameters by removing those revealing correlations. Subsequent studies where this methodology was employed outperformed other contemporary methods, including the eigenvalue-based method and the principal components analyisis[47, 48].

To further enhance and validate the estimability results, an analysis utilizing the eigenvalues of the Hessian matrix can be conducted to obtain mathematical evidence that a global optimum can be found within the feasible region defined by the selected estimable parameters[45].

2.6.4 Data Reconciliation

To ensure accurate characterization of the observed data, the application of data reconciliation can be considered. Data reconciliation is a procedure that adjusts both the measured data and the obtained parameters so that the yielded values align with conservation laws and other constraints. This adjustment is necessary since process variables are prone to inherit measurement errors and variabilities. These errors can stem from random or systematic sources. Systematic errors are linked to malfunction, poor conditions, and calibration issues, thus arising from non-random events. Random errors, on the other hand, result from fluctuations in power, environmental conditions, or network transmissions, making them unavoidable. While various techniques may be employed, the general data reconciliation problem can be addressed by the weighted least-squares minimization of the error between the measurements and the model predictions[49, 50]. Data reconciliation has been successfully

applied to various industrial processes. Additionally, it has been employed in smaller-scale studies, showcasing versatility and applicability.

2.6.5 Uncertainty Analysis

In the pursuit of mitigating measurement errors, incorporating uncertainty analysis becomes essential when dealing with parameter and variable measurements prone to errors. Uncertainty studies can be conducted to enhance equipment behaviour prediction and facilitate higher-quality risk analysis. However, performing such studies necessitates a comprehensive knowledge of all the factors contributing to uncertainty.

A useful resource in uncertainty analysis is the Guide to the Expression of Uncertainty in Measurement (GUM). According to this guide, uncertainty refers to the "doubt about the validity of the result" and emphasizes that a measurement is only considered complete whenever it is accompanied by its corresponding uncertainty. GUM showcases two types of uncertainty evaluation: Type A, which involves evaluation through statistical analysis, and Type B, where the evaluation encompasses means other than statistical analysis. The guide delves into relevant topics such as standard uncertainty, combined uncertainty, and propagation of uncertainty, thereby revealing itself as an important tool for carrying out complete sensitivity analyses[51, 52].

2.7 Literature Review

Previous works with similar objectives and methodologies to those discussed in this dissertation have been published. The following examples report some of these works. Suzuki *et al.* employed parameter estimation in a SMB in order to enhance the mathematical model using operational data collected from a plant involved in the production of high fructose corn syrup. By employing Tikhonov regularization and conducting clustering analysis, the authors revealed that using the operational data, the models could indeed be improved and be made to yield more accurate predictions. However, it was observed that some parameters that described the adsorption equilibrium were especially challenging to predict due to correlations or low sensitivity[53]. In another similar work, Santos *et al.* developed a parameter estimation methodology, coupled with uncertainty analysis, that uses PSO to overcome the issue of local minima and enable the desired sensitivity analysis. The authors validated the model by integrating system-in-the-loop technology and finally using data present in the open literature, showcasing that the methodology developed is reliable for parameter estimation in SMB chromatography[54]. Additionally, other studies have reported using PSO, or similar versions of the method, to perform parameter estimation in an array of distinct problems[55-59]. These studies serve as vital references, and this dissertation aims to build upon their insights.

3 Materials and Methods

3.1 Materials

The underlying thesis for this dissertation, authored by Albertina Rios, used three distinct materials, Sephadex G-50, Q Sepharose FF, and Hydroxyapatite, to accomplish the separation of two model proteins, Bovine Serum Albumin (BSA) and Myoglobin (Mb)[4]. Within this study, the focus was solely placed on one of these materials, Sephadex G-50.

Myoglobin, a member of the globin protein family, is mainly responsible for the binding and transport of oxygen in most vertebrates due to the presence of a heme group. This protein is mostly present in cardiac and skeletal muscles. Myoglobin has a mass of approximately 17800 Da[60, 61]. Figure 5 presents a three-dimensional representation of this protein's structure.

Bovine Serum Albumin, abbreviated as BSA, is a specific type of albumin derived from bovine sources and stands as the most abundant protein in blood plasma. It is characterized by its smaller size, in relation to other blood proteins, and its high solubility and stability[62]. The BSA molecule consists of 563 amino acids, which translates to a substantial mass of 66400 Da, surpassing myoglobin. This difference permits using sieve chromatography (size exclusion) to separate these proteins[63]. Figure 5 visually depicts BSA's three-dimensional structure, distinctly showing the increase in size compared to the structure of myoglobin.

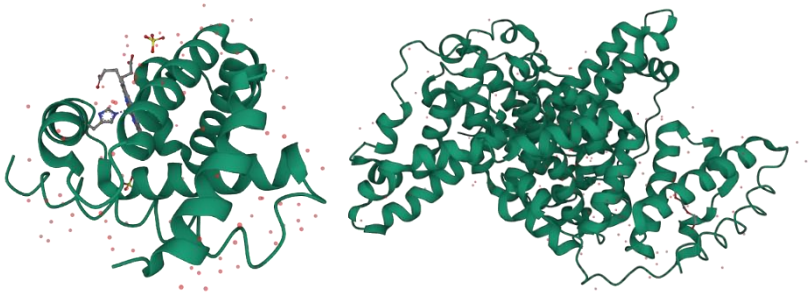

Figure 5 – Three-Dimensional models of the myoglobin (left) and BSA (right) proteins. Images obtained from RCSB PBD [64, 65]

The material studied, Sephadex G-50, is an organic size exclusion gel with a sieve range spanning from 1500 to 30000 Da, meaning that myoglobin may enter the gel pores, whereas BSA can freely traverse the gel due to its higher molecular mass[66].

The software employed in this work comprised MATLAB® R2022b Update 4 (9.13.0.2166757), MATLAB® R2023a Update 1 (9.14.0.2239454), and gPROMS 7.0.7.55178. Both versions of MATLAB® were utilized to develop the necessary code scripts, build the objective functions, and execute the PSO algorithm. The software gPROMS was used to write the code for the

mathematical models that describe the SMB operation, and the adsorption mechanisms, to generate the simulated data. These software tools were used in combination during the parameter estimation process. Detailed explanations will be provided in further sub-sections.

Throughout this work, three different computers were used. A desktop computer equipped with an Intel® Core™ i5-7400 CPU @ 3.00GHz processor and 4GB of RAM was used to modify the gPROMS models into a parameter estimation setting and to develop the MATLAB® scripts encompassing the objective functions and preliminary tests. Subsequently, these files were sent to a more powerful desktop computer featuring an Intel® Core™ i9-9900K CPU @ 3.60GHz processor and 32 GB of RAM to perform the optimization with a higher number of simulations, thereby enhancing result quality while decreasing the time required in a single simulation. Additionally, a portable computer powered by an AMD Ryzen 5 3550H Processor and 8 GB of RAM was utilized to develop certain MATLAB® scripts, whenever the other computers were occupied with parameter estimation simulations.

3.2 Methodology, Methods and Models

The methodology considered for the development of this master's dissertation had the following chronological structure:

1. Modification of the existing gPROMS mathematical models, which describe protein separation, into models ready for parameter estimation.
2. Development of the MATLAB® scripts that facilitate parameter estimation and creation of the scripts containing the objective functions.
3. Execution of parameter estimation for the fixed-bed experiments, using Particle Swarm Optimization with a substantial number of iterations and particles, to obtain reference parameters for subsequent validation of results and create particle density regions for further investigation.
4. Elaboration of an estimability analysis to determine the correlation between chosen estimation parameters and quantify their influence on the final outcome.
5. If necessary, remove parameters from further estimations and adapt their lower and upper bounds based on the particle density regions and estimability analysis results. This step aims for the simplification of the total parameter estimation.
6. Execution of total parameter estimation, using the fixed-bed experiments, and the SMB experiment, to identify a set of parameters that would create a mathematical model capable of accurately fitting the experimental data from all the experiments.
7. Carry out an uncertainty analysis by computing each parameter uncertainty, building confidence regions, and propagating the uncertainties to the final result.

Although the methodology underwent modifications from its initial form as designed at the beginning of the dissertation, this final version will, nonetheless, attempt to fulfill all the proposed objectives for this dissertation.

It was considered that the adsorption isotherm assumed by Albertina Rios for the Sephadex G-50 separation, a linear isotherm, was the most suitable for the obtained separation data since size exclusion separation, that occurs within the Sephadex G-50 matrix, is a straightforward process, where the retention time of a component in the column is determined solely by its size. Additionally, no competition or site inhibition occurs during this separation.

It is worth noting that different adsorption materials possess different models due to the distinct underlying adsorption mechanisms governing the components' adsorption in each matrix. As this dissertation focus on only one material, only two of these models were investigated and modified: the fixed bed model and the global SMB model for Sephadex G-50.

To prepare these models for parameter estimation, the desired parameters to be studied must be set as variables beforehand within the gPROMS software. This requires meticulous adjustments to the gPROMS models' code. Furthermore, the connection between gPROMS and MATLAB® must be built by creating several gPROMS tasks and developing MATLAB® code. This connection will facilitate data exchange between the two software and provide the means necessary to perform gPROMS simulations using values supplied by MATLAB®. Detailed explanations of the studied models are provided in the following sub-sections.

3.2.1 Fixed Bed Modelling

As previously mentioned, Sephadex G-50 is a size-exclusion gel. As a result, the separation is achieved due to the difference in molecular size/molecular weight. Notably, within this gel, effective adsorption does not occur, as elution time depends on the molecules' pore accessibility. Molecules with greater pore accessibility reveal slower velocity through the column. Hence, the underlying mechanism is consequently described by a simpler model, requiring fewer parameters to be completely defined.

The model employed was considered isothermal, assuming a constant mobile phase velocity (diluted mixtures) and the absence of interactions between solutes. The gel's particles were considered spherical and uniform. Additionally, plug-flow with axial dispersion was considered for the movement of the fluid phase[66]. With these assumptions, the mass balance to the fluid phase, considered for this model, is represented by equation 3.1:

$$\frac{\partial C_i}{\partial t} = D_{ax}\frac{\partial^2 C_i}{\partial z^2} - u_i\frac{\partial C_i}{\partial z} - \frac{(1-\varepsilon_b)}{\varepsilon_b}\frac{\partial \bar{q}_i}{\partial t} \qquad (3.1)$$

In this equation, the parameter D_{ax} refers to the axial dispersion, u_i is the interstitial velocity, ε_b is the fixed bed porosity, C_i is the concentration in the fluid phase of component i and $\bar{q}_i$ is the average adsorbed concentration of component i. Lastly, z represents the axial direction.

The mass balance to the particle is described by the linear driving force (LDF) equation:

$$\frac{\partial \bar{q}_i}{\partial t} = k_h(q_i^* - \bar{q}_i) \tag{3.2}$$

in which q_i^* is the adsorbed concentration in equilibrium with the fluid phase concentration of component i, while k_h is the mass transfer coefficient.

Lastly, the equilibrium relation between the liquid and solid phases can be represented by the following expression 3.3, equivalent to a linear isotherm equation:

$$q_i^* = K_{SEC} C_i \tag{3.3}$$

In this equation, K_{SEC} represents the size exclusion constant.

To correctly simulate the model, the following initial and boundary conditions were considered:

Initial Conditions:

$$t = 0 \rightarrow C_i = \bar{q}_i = 0 \tag{3.4}$$

Boundary Conditions:

$$z = 0 \rightarrow D_{ax} \frac{\partial C_i}{\partial z} = u_i\left(C_i - C_i^F\right) \tag{3.5}$$

$$z = L \rightarrow \frac{\partial C_i}{\partial z} = 0 \tag{3.6}$$

with L being the column length and C_i^F being the concentration of component i at the feed inlet.

The parameters that were taken into account for estimation in this study were: ε_b, D_{ax}, Q_{feed}, k_h, and K_{SEC}. The parameter Q_{feed} represents the feed flow for the fixed-bed column. The column length was considered to be constant and known. Although ε_b and Q_{feed} could be readily determined through experimental work and precise measurements, respectively, they were selected for estimation to understand their uncertainty. The parameter u_i, on the other hand, was not considered for estimation as it can be derived through the feed flow Q_{feed} using expression 3.7:

$$u_i = \frac{Q_{feed}}{A\,\varepsilon_b} \tag{3.7}$$

where A is the known area of the column cross-section.

3.2.2 SMB Modelling

To develop a global SMB model, it is possible to utilize the already available fixed bed model, and incorporate the necessary mass balance to the nodes. Furthermore, the model should include the necessary global balances and represent the structure of the additional equipment involved, such as pipes and manifolds.

The global balances are independent of the material used and can be described by the following expressions:

$$Q_I = Q_D + Q_{IV} \tag{3.8a}$$

$$Q_{II} = Q_I - Q_E \tag{3.8b}$$

$$Q_{III} = Q_{II} + Q_F \tag{3.8c}$$

$$Q_{IV} = Q_{III} - Q_R \tag{3.8d}$$

where *I*, *II*, *III*, and *IV* represent the section of the SMB equipment, and the subscripts *D*, *E*, *F*, and *R*, are related to the desorbent, extract, feed, and raffinate streams, respectively.

From 3.8a to 3.8d, the following mass balance can be obtained:

$$Q_D + Q_F = Q_E + Q_R \tag{3.9}$$

The following expressions 3.10a to 3.10e, presented below, are equations that accurately depict the dynamics of a simulated moving bed equipment at the inlet/outlet nodes or between columns.

Desorbent node: $$C_{i,I}^{in} = \frac{Q_{IV}\, C_{i,IV}^{out}}{Q_I} \tag{3.10a}$$

Extract node: $$C_{i,II}^{in} = C_{i,I}^{out} \tag{3.10b}$$

Feed node: $$C_{i,III}^{in} = \frac{(Q_{II}\, C_{i,II}^{out} + Q_F\, C_{i,F})}{Q_{III}} \tag{3.10c}$$

Raffinate node: $$C_{i,IV}^{in} = C_{i,III}^{out} \tag{3.10d}$$

Between columns: $$C_{i,k}^{in} = C_{i,k-1}^{out} \tag{3.10e}$$

To account for the transport of material between columns or before/after the inlet/outlet ports, transfer lines were introduced into the SMB model, in which plug-flow is assumed; therefore, dead volumes of the system were also considered.

The SMB operational flowrates, Q_F, Q_D, Q_E, and Q_{IV} were investigated during total parameter estimation. Additionally, each column's bed porosity and Peclet Number were chosen as estimable parameters; therefore, total estimation consisted in 21 distinct parameters.

3.2.3 Particle Swarm Optimization Algorithm

To ensure the correct application of a PSO algorithm, it is imperative to select an appropriate number of particles (N_p) and number of iterations (N_{iter}) to enable the convergence of the method without compromising excessive time and/or computational power. Furthermore, it is essential to choose suitable values for the other PSO parameters, such as the inertia weight (ω_O) and the learning factors (C_f). In the PSO algorithm employed in this study, these parameter values were predefined by its developers and were deemed reasonable for the intended purpose. Additionally, the number of decisions variables (N_{DV}), or parameters, and the values for the lower and upper position bounds (x_{min} and x_{max}, respectively) must also be properly established to construct a feasible region for the optimization problem.

The particle swarm optimization algorithm diligently follows a sequence to find an optimum within the region defined by the problem's decision variables. Figure 6 provides a simplified representation of the PSO algorithm's structure applied in this study.

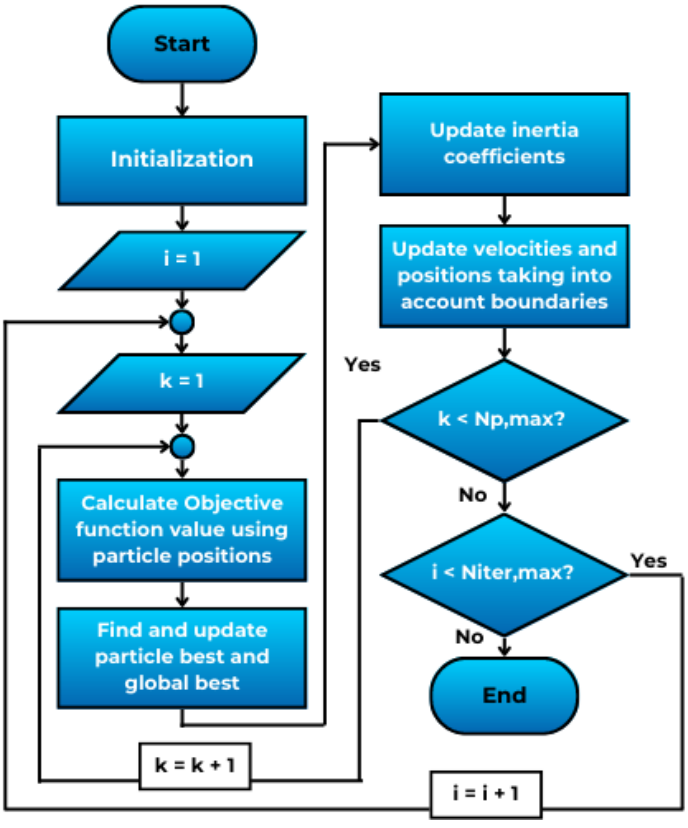

Figure 6 - PSO algorithm employed in this work

The particle swarm optimization algorithm applied in this dissertation starts with an initialization step, where a position x is randomly given to each particle using equation 3.11:

$$x_{i,k,n} = x_{min}(n) + r^{*}\left(x_{max}(n) - x_{min}(n)\right) \quad (3.11)$$

in which i represents the iteration value, k shows the particle number, n denotes the decision variable and r^* is a random number between 0 and 1. The maximum velocity achievable by the particles in this optimization is also defined in this step and its equivalent to the upper limit of the parameter in question ($v_{max} = x_{max}$).

Once the positions of each particle have been determined, the next step involves calculating the objective function value. Subsequently, the best position of an individual particle, P^{best}, and the overall best position for the swarm up to the current iteration, G^{best}, are updated. The determination of the best positions is based on the minimum objective function value yielded by the particles.

The next step encompasses a value update for the inertia weight parameters. These parameters will evolve according to expression 3.12:

$$\omega_{O,j}^{*} = \omega_{O,j} + \frac{\left(C_{f,j} - \omega_{O,j}\right) * i}{N_{iter}} \tag{3.12}$$

where, once again, i represents the iteration and j = 1,2. $\omega_{O,j}^{*}$ is the updated value of the inertia weight to be used in the current iteration, while $\omega_{O,j}$ is the value utilized in the previous iteration. In the first iteration, *i.e.*, $i = 1$, $\omega_{O,j}^{*} = \omega_{O,j}$.

Following the inertia weight update, velocities are obtained using the following equation:

$$v_{i,k,n} = \omega_{O,1}^{*} r_1\left(x_i^{G_{best}} - x_{i-1,k,n}\right) + \omega_{O,2}^{*} r_2\left(x_i^{P_{best}} - x_{i-1,k,n}\right) \tag{3.13}$$

where r_1 and r_2 are random numbers between 0 and 1. It is important to note that the following restriction $v_{i,k,n} \leq v_{max}$ must be obliged. If the constraint is violated, the new velocity will be set to the maximum velocity value, ensuring that the restriction is once again complied.

The particle position is then updated using the following equation:

$$x_{i,k,n}^{*} = v_{i,k,n} + x_{i-1,k,n} \tag{3.14}$$

The new position must not compromise the region's limits, which means $x_{min} \leq x_{i,k,n}^{*} \leq x_{max}$. Once again, in the event that this restriction is violated, the new positions will assume boundary values, *i.e.*, $x_{i,k,n}^{*} = x_{min}$ or $x_{i,k,n}^{*} = x_{max}$ depending on the type of boundary breach. If the restriction is not broken, then $x_{i,k,n}^{*}$ is the particle position for the next iteration.

Lastly, after the aforementioned step, if k is lower than the total number of particles, N_p, then the algorithm is performed again until all particles have been analysed. Once this is concluded if i is less than the chosen number of iterations, N_{iter}, then all the steps explored above are repeated until the designated number of iterations has been reached.

This algorithm will generate important insights, such as the minimum objective function value obtained throughout this process, as well as identifying the particle, *i.e.*, which values for the decision variables, achieved the best solution. The position of each particle and the respective objective function values are archived to be applied in further studies. In this study, the main outcomes obtained from the PSO algorithm are the estimated parameter values.

3.2.4 Objective Function

The objective functions to be minimized by the PSO algorithm were designed to guide the PSO algorithm in searching for a set of parameters that would minimize the residuals between the generated mathematical model and the experimental points.

To achieve this goal, the MATLAB® script that defined the objective function had the following structure: Each experiment used for the parameter estimation was simulated using the "gOMATLAB" command, which allowed MATLAB® to call gPROMS to perform the desired simulations. As mentioned before, the parameters used in the simulations were determined by the positions of each decision variable ($x_{i,k,n}$), obtained through the PSO algorithm. This resulted in a matrix with the simulated data for each experiment. The experimental data collected by Albertina Rios was imported to MATLAB®, and organized in a predefined matrix. This process was repeated for all relevant experiments that would aid the parameter estimation. In the case of SMB experiments, an additional calculation was necessary to determine the average concentration of each cycle at the extract/raffinate port. Since the available experimental data only provided this information, while the simulated data yielded concentration history at the outlets, this calculation was necessary to calculate the residuals accurately. Therefore, in studies involving SMB experiments, a section of the MATLAB® script is dedicated to this step. Lastly, the residuals between the experimental data and the simulated data, stored in matrices, were then calculated using the Mean Squared Error (MSE), computed using the following expression:

$$\frac{1}{N}\sum_{i=1}^{N}\left(y_{exp} - y_{sim}\right)^2 \tag{3.15}$$

In this equation, N is the number of times this discrepancy is calculated, which coincides with the number of experimental points, y_{sim} represents the simulated data, and y_{exp} represents the experimental data.

The sum of the residuals of all the experiments is the value to be minimized by the PSO. Due to the distinct number of experimental data among experiments, weighing factors were applied to ensure the equal influence of each experiment in the residual calculation.

The structure is summarized in the following figure:

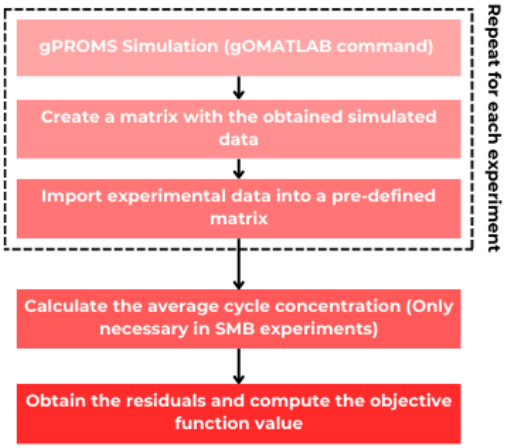

Figure 7 - Summarized structure of the developed objective functions

It is important to emphasize that when comparing objective function values, it is imperative only to make the comparison within experiments. Specifically, only the parameters obtained by Rios and the estimated parameters for the same experiment should be compared. This is because a higher number of experimental points, among other factors, can result in a higher objective function value, even if the resulting mathematical model better fits the experimental data. This occurrence is due to the nature of the objective function formulation and was anticipated and accounted for.

3.2.5 Estimability Analysis

The estimability analysis employed in this dissertation is based on the Gram-Schmidt principle of orthogonalization, depicted in Figure 8. Each parameter, S_i, has an assigned sensitivity vector, S_i'. The selection process involves identifying the parameter with the highest sensitivity, after which the remaining vectors are represented orthogonally in the plane perpendicular to the chosen parameter's sensitivity vector. This will be iterated until all parameters are selected or until the sensitivity of the largest parameter falls below a predefined cut-off value[47].

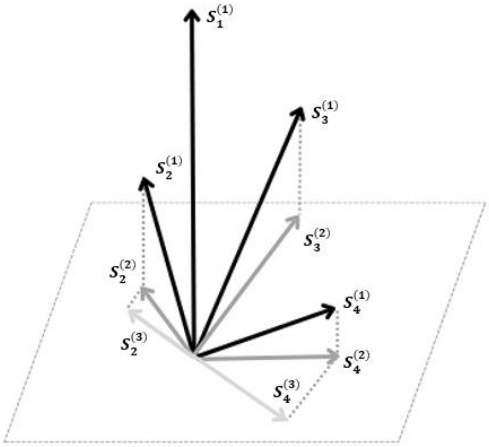

Figure 8 - Illustration of the Gram-Schmidt orthogonalization. Adapted from [47]

Using this methodology, the initial step is to create the sensitivity matrix.

Considering the following expression:

$$y = f(x(t), \theta, t) \tag{3.16}$$

in which y is the vector that holds the outputs, $x(t)$ the vector with the state variables values, θ the vector with the model parameters, and t is the time, we can determine the sensitivity matrix coefficients using equation 3.17:

$$S'_{i,p} = \frac{\theta_p}{y_i} \frac{y(t, \theta_p + \Delta\theta_p) - y(t, \theta_p)}{\Delta\theta_p}\Bigg|_{t_n} \tag{3.17}$$

with i = 1,2, 3, …, N_y; p = 1,2, 3, …, N_p; n = 1,2, 3, …, N_t

and where y_i is the i-th output obtained through the model, θ_p, is the parameter p, t_n is the time instant evaluated, N_y is the number of output variables, N_p is the number of parameters, and N_t is the number of chosen time-steps. These calculations will result in a sensitivity matrix S, that demonstrates parameter influence in its columns:

$$S = \begin{vmatrix} S'_{1,1}\big|_{t_1} & \cdots & S'_{1,n_p}\big|_{t_1} \\ \vdots & \ddots & \vdots \\ S'_{1,1}\big|_{t_n} & \cdots & S'_{1,n_p}\big|_{t_n} \\ \vdots & \vdots & \vdots \\ S'_{n_y,1}\big|_{t_1} & \cdots & S'_{n_y,n_p}\big|_{t_1} \\ \vdots & \ddots & \vdots \\ S'_{n_y,1}\big|_{t_n} & \cdots & S'_{n_y,n_p}\big|_{t_n} \end{vmatrix}_{n_y - n_t \times n_p} \tag{3.18}$$

Notably, the coefficients are normalized because the parameters can have different orders of magnitude.

A computation of the magnitude of each column (each parameter) has then proceeded using equation 3.19:

$$M(p) = S_p^T S_p \tag{3.19}$$

In the expression above, S_p is the column p of the sensitivity matrix, while T denotes the transpose of the associated vector.

Through equation 3.19, it is possible to know the column, and therefore parameter, that has the biggest magnitude and consequently, the highest influence in the final result, S_{max}. It is now necessary to orthogonalize matrix S in relation to this column using expression 3.20:

$$S_0 = S_{max}(S_{max}^T S_{max})^{-1} S_{max}^T S \tag{3.20}$$

The residual matrix, R, is then computed using both matrices S and S_0:

$$R = S - S_0 \tag{3.21}$$

Upon obtaining matrix R, the magnitude of its columns is subsequently calculated, and the next parameter is chosen as the reference for orthogonalization. As mentioned, this iterative process continues until all parameters have been analysed or until the largest magnitude attained falls below a predefined tolerance value.

The algorithm for this process is illustrated in Figure 9:

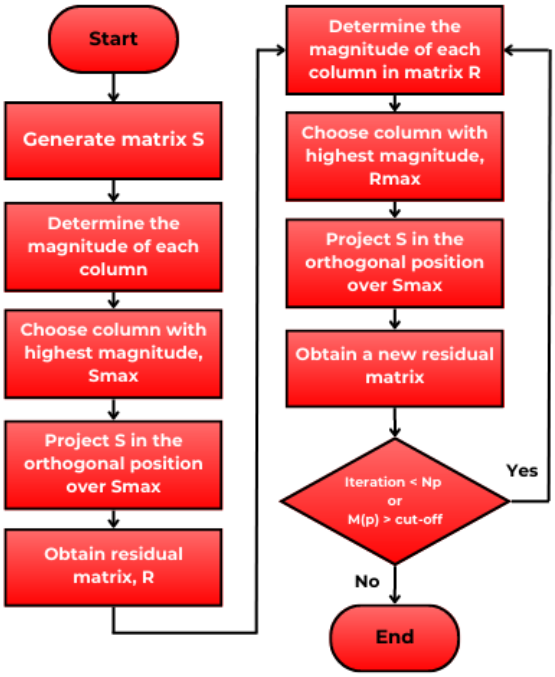

Figure 9 - Estimability Analysis Algorithm employed in this work

This analysis enables the removal of correlated parameters, leading to a simplification of the optimization problem. However, a subsequent analysis involving the Hessian matrix is required to ensure mathematical certainty regarding the selected estimable parameters' ability to yield a global optimum.

The Hessian matrix, denoted as H, can be derived through the coefficients of the matrix S_f. This matrix is similar to the sensitivity matrix S; however, it exclusively includes the parameters identified as estimable through the orthogonalization methodology. Thus, the Hessian matrix is obtained through the following equation 3.22:

$$H = 2S_f^T S_f \tag{3.22}$$

Since S_f reveals the region created by the estimable parameters, then H will therefore indicate if a global optimum exists. A positive-definite Hessian matrix confirms that the region defined

by the selected parameters will yield a global minimum. Otherwise, the absence of a positive-definite Hessian matrix implies a lack of mathematical proof for convergence to the minimum optimum or even the existence of such optimum within the parameter-defined region.

This analysis was implemented in a MATLAB® script and performed in the fixed-bed experiments for myoglobin, BSA, and a binary mixture. The aim was to identify correlations and understand the influence of individual parameters in the final result, thereby simplifying the total parameter estimation. Additionally, a cut-off value was introduced so parameters with negligible impact could be discarded from the estimation process. The considered cut-off value was set at 100.

3.2.6 Uncertainty Evaluation

The uncertainty of the parameters was evaluated based on the outcomes yielded by the total estimation. Throughout the Particle Swarm Optimization, all particle positions, $x_{i,k,n}$, were recorded and stored. However, only those who fulfil the Fisher-Snedecor test were utilized in the uncertainty evaluation. To quantify this uncertainty, u_p, the following expression was utilized:

$$u_p = K \times \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y^{PSO} - \bar{y})^2} \tag{3.23}$$

in which $\bar{y}$ can be defined by expression 3.24, K is the coverage factor, which was assumed to be 2 (approximately 95% of confidence), N is the total number of values obtained for each parameter, i denotes a particle number, and y^{PSO} are the values yielded by the PSO that fulfilled the Fisher-Snedecor test for a certain parameter.

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y^{PSO} \tag{3.24}$$

Having obtained the uncertainty value for each parameter, the design of confidence regions became feasible. While not the primary focus of this dissertation, they provide useful insight into the behaviour of parameter uncertainty. Furthermore, they form an integral part of the uncertainty evaluation, a key objective of this work.

The propagation of the uncertainty into the final result can also be determined by utilizing equation 3.25:

$$U_c = \sqrt{\frac{\sum_{i=1}^{p} {u_{pi}}^2}{\left(\sum_{i=1}^{p} X_i\right)^2}} \tag{3.25}$$

in which X_i is a vector with the best values for each parameter yielded by the PSO.

4 Results and discussion

4.1 Fixed-Bed Experiments

In order to effectively study and determine the parameters for the separation process under consideration, first, an analysis was performed on independent fixed bed experiments. This analysis entailed the implementation of a Particle Swarm Optimization with a significant number of particles and iterations. These values were selected with a balance between computer resources and analysis time in mind, all the while enabling the acquisition of a set of parameters that would create a mathematical model with perfect alignment with the experimental data from the fixed bed experiments. These estimated parameters yielded in this step will serve as a reference for further estimations.

For the specific fixed bed experiment focused solely on the myoglobin protein, the estimated parameters were compiled in Table 1, alongside the parameters of the mathematical model used by Rios. The estimation process employed 100 iterations and 100 particles. Additionally, the value of the objective function can be observed in this table. Figure 10 presents a comparison between the experimental data points and the mathematical model computed with Rios' parameters (left) and with the estimated parameters (right).

While the graphical representation may not reveal noticeable differences due to the relatively low nature of the objective function value, through Table 1, it is possible to highlight a significant disparity between the axial dispersion coefficient, D_{ax}, values. This suggests that the axial dispersion may be slightly lower than the previously assumed value.

Table 1 – Rios' and Estimated set of parameters for the Mb experiment

	Rios' Parameters	Estimated Parameters
ε_b	0.390	0.395
D_{ax} $(m^2 \cdot s^{-1})$	3.83 x 10^{-7}	3.28 x 10^{-7}
Q_{feed} $(m^3 \cdot s^{-1})$	3.17 x 10^{-8}	3.25 x 10^{-8}
K_{SEC}	0.300	0.333
k_h (s^{-1})	0.175	0.999
f_{obj}	1.23 x 10^{-5}	5.01 x 10^{-6}

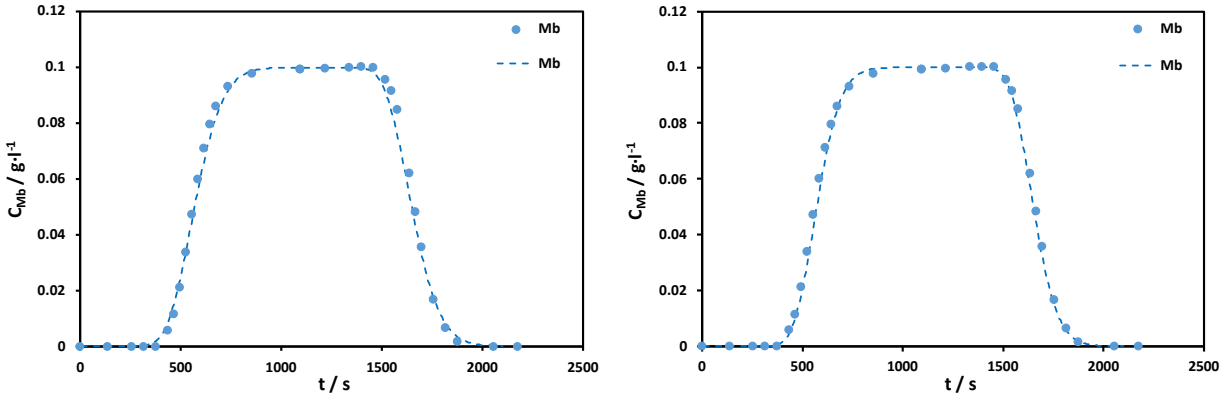

Figure 10 - Graphical comparison between Mb experimental data (points) and the models (line) with Rios' parameters (left) and the estimated parameters (right)

The mass transfer coefficient, k_h, exhibited a noticeable variation; however, this deviation was suspected of being misleading, as the estimated value closely approached the predefined upper limit of the parameter (1.0). Moreover, no significant changes were observed in the resulting mathematical model. Hence, it was inferred that this particular parameter may have minimal influence on the outcome. The size exclusion coefficient, K_{SEC}, demonstrates an increase in value, still within the expected range. This suggests that myoglobin has a greater degree of pore accessibility than previously determined. As for the remaining parameters, only minor deviations were observed. These deviations collectively contribute to a decrease in the objective function value, which confirms that the estimated parameters create a model with better alignment to the experimental data than the model generated with Rios' parameters.

The procedure presented beforehand was also employed in the BSA fixed bed experiment, employing the same number of particles and iterations. Table 2 compiles the parameters and objective functions' values, while Figure 11 presents a graphical comparison.

Table 2 – Rios' and Estimated set of parameters for the BSA experiment

	Rios' Parameters	Estimated Parameters
$\boldsymbol{\varepsilon_b}$	0.39	0.391
$\boldsymbol{D_{ax}\ (m^2 \cdot s^{-1})}$	3.83 x 10^{-7}	5.40 x 10^{-8}
$\boldsymbol{Q_{feed}\ (m^3 \cdot s^{-1})}$	3.17 x 10^{-8}	3.16 x 10^{-8}
$\boldsymbol{K_{SEC}}$	0	0.028
$\boldsymbol{k_h\ (s^{-1})}$	0.0564	0.03018
$\boldsymbol{f_{obj}}$	1.46 x 10^{-1}	3.26 x 10^{-2}

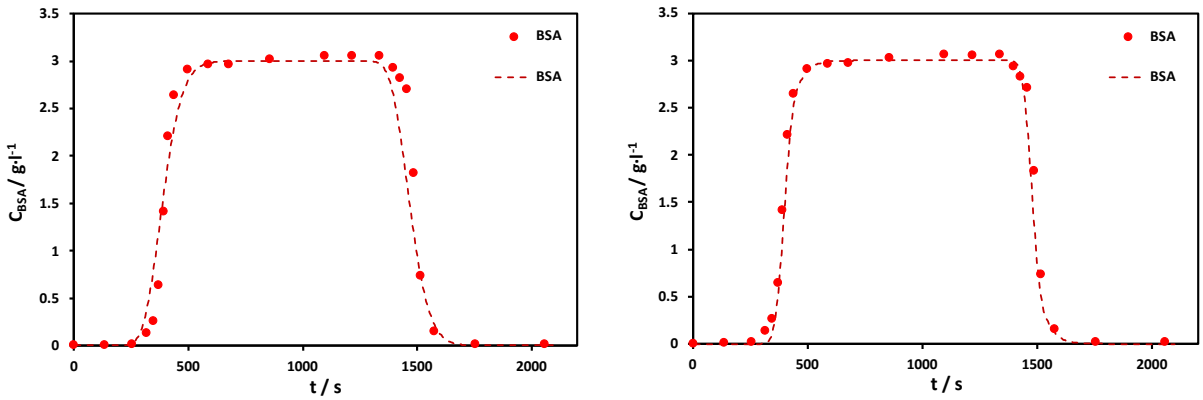

Figure 11 - Graphical comparison between BSA experimental data (points) and the models (line) with Rios' parameters (left) and the estimated parameters (right)

In this experiment, contrasting the myoglobin experiment, slight differences can be observed in the graphs of the mathematical models, especially in terms of curvature, resulting in a better fit to the experimental data. This curvature is a direct consequence of axial dispersion, suggesting a disparity between the estimated value for this parameter and the previously determined value. The values presented in Table 2 confirm this variation. It is worth noting that the axial dispersion is once again lower than the initial value, albeit with a larger deviation than the myoglobin experiment. The estimated mass transfer coefficient resembles the one derived from Rios' calculations.

An interesting change is the value of the size exclusion constant parameter for BSA, which has shifted from 0 to 0.028. Originally, this protein was expected to exhibit no interaction with the gel due to its size surpassing the material's sieve dimensions. However, this variation indicates a certain level of influence exerted by the gel pores in the BSA protein. Nonetheless, when comparing it to the value of the size exclusion constant for myoglobin, this variation remains negligible, as it does not account for 10% of the corresponding myoglobin parameter.

Once again, ε_b and Q_{feed} did not reveal any notable differences that would affect the separation mechanism, though ε_b reported a slightly lower value compared to the one yielded at the myoglobin experiment.

Finally, the parameters of a binary mixture of these two proteins were studied. Two distinct sets of experimental points were considered for this fixed bed experiment, in which different concentrations of myoglobin and BSA were utilized. In the first experiment (Exp. 1), BSA was present at a concentration of 1.75 g·l^{-1} while 0.5 g·l^{-1} was used for Mb. In the second experiment (Exp. 2), the concentrations for BSA and Mb were 3.0 g·l^{-1} and 0.1 g·l^{-1}, respectively. The parameters obtained using 100 iterations and 100 particles and Rios' parameters, are presented in Table 3. Their corresponding objective function values are also included. Figure 12 and Figure

13 illustrate a graphical comparison between both sets of experimental data points and the mathematical models based on Rios' parameters (left) and the estimated parameters (right).

Table 3 – Rios' and Estimated set of parameters for the binary experiments

	Rios' Parameters	Estimated Parameters
ε_b	0.39	0.395
$D_{ax}\ (m^2 \cdot s^{-1})$	3.83 x 10^{-7}	9.74 x 10^{-8}
$Q_{feed}\ (m^3 \cdot s^{-1})$	3.17 x 10^{-8}	3.17 x 10^{-8}
$K_{SEC,BSA}$	0	0.015
$K_{SEC,Mb}$	0.3	0.298
$k_{h,BSA}\ (s^{-1})$	5.64 x 10^{-2}	0.566
$k_{h,Mb}\ (s^{-1})$	0.175	0.047
f_{obj}	0.0151	0.0107

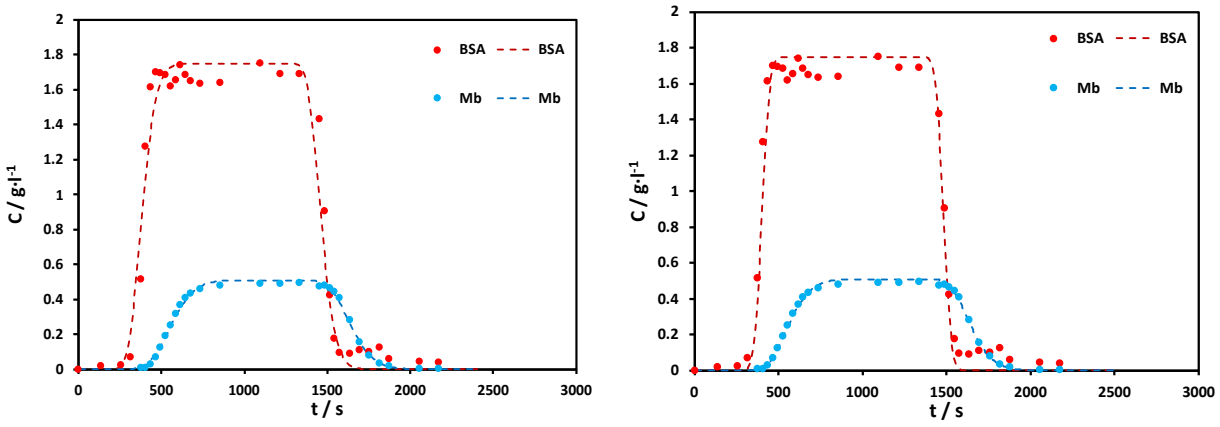

Figure 12 - Graphical comparison between the binary experimental data (points) and the models (line) with Rios' parameters (left) and the estimated parameters (right) – Exp. 1

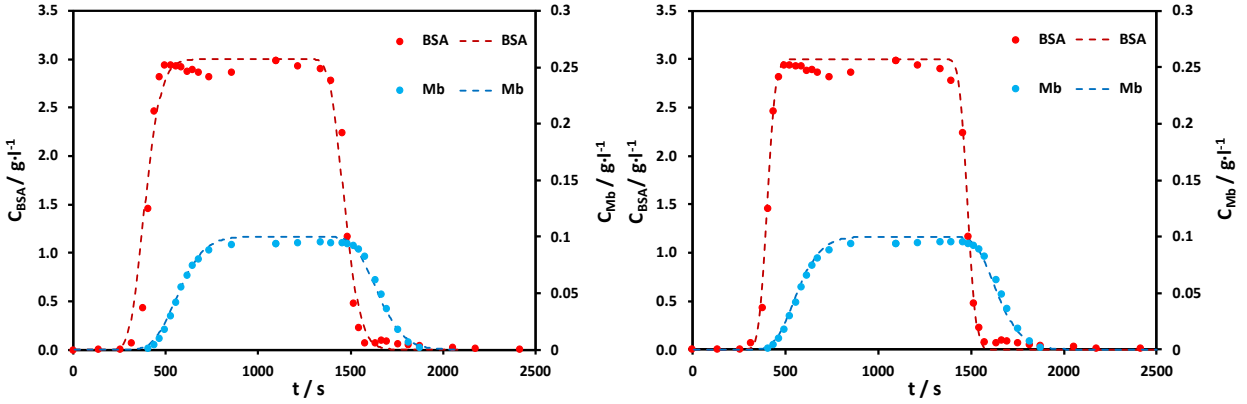

Figure 13 - Graphical comparison between the binary experimental data (points) and the models (line) with Rios' parameters (left) and the estimated parameters (right) – Exp. 2

Table 3 reveals that the axial dispersion parameter remains consistently smaller than the value calculated by Rios. In the case of the binary mixture, the obtained value falls within the range between the values obtained for the individual protein studies. The value for the size exclusion constant of BSA, $K_{SEC,BSA}$, exhibits a slight deviation from 0, which was expected considering the results of the last experiment. Conversely, $K_{SEC,Mb}$ reveals no significant change. The mass transfer coefficients reveal values that are very dissimilar to the ones yielded by Rios' calculations. This revelation further supports the suspicion that these parameters have minimal influence on the final result.

To verify this intuition, external analysis, such as estimability analysis or examining specific parameter values and analysing their corresponding objective functions can be performed.

Images representing particle density regions, which relate two parameters and the yielded objective functions, were obtained from the fixed-bed parameter estimation. These images allow for several conclusions to be drawn. They provide insights into the need to change each parameter lower and upper bounds for total parameter estimation. Additionally, it is possible to obtain some preliminary information regarding the influence of the parameter in the objective function value, thereby validating the hypothesis that certain parameters, such as the mass transfer coefficients, exert minimal influence on the resulting mathematical model. All the particle density regions generated during the fixed bed parameter estimations are presented in Appendix A.

Indeed, these images highlight that the parameters, $k_{h,BSA}$ and $k_{h,Mb}$, were able to yield very low objective function values regardless of the parameter value within their designated bounds. This observation is particularly evident in the provided examples illustrated in Figure 14.

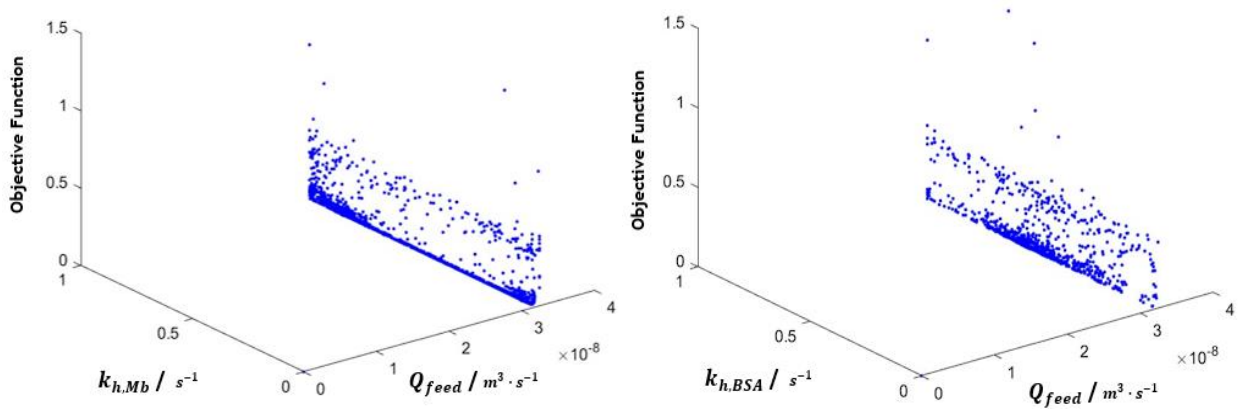

Figure 14 - Particle density regions for the parameters $k_{h,Mb}$ (left) and $k_{h,BSA}$ (right), with Q_{feed} as reference, compared to the objective function value yielded

Hence, this observation suggests that estimating these parameters would lack significance, as a wide range of values could produce nearly the same results.

From the analysis of the images pertaining to the feed flow of the fixed bed, Q_{feed}, including the aforementioned examples in Figure 14 and the left graph of Figure 15, it was deduced that the boundaries defined for this parameter might have been too restrictive. As a result, the lower and upper limits for this parameter were slightly expanded.

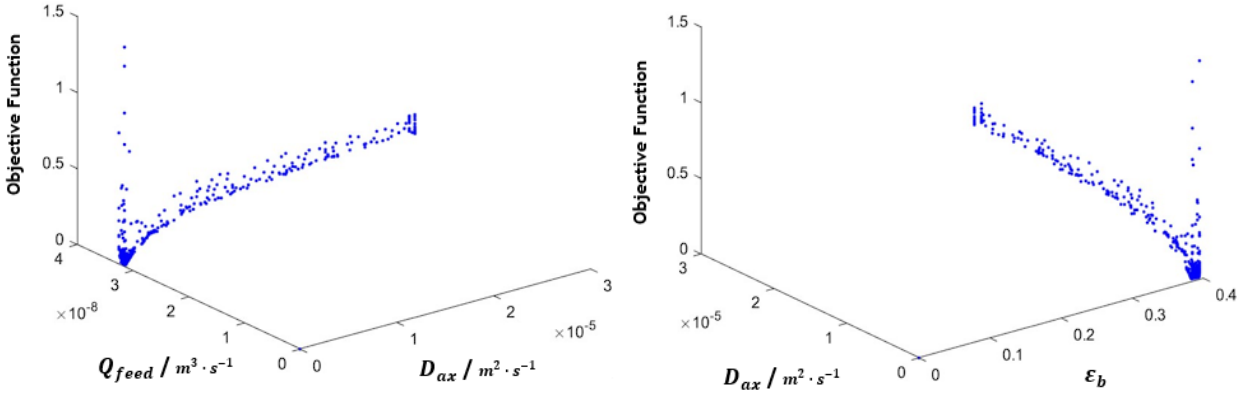

Figure 15 - Particle density regions for the parameter D_{ax}, with Q_{feed} (left) and e_b (right), compared to the objective function value yielded

The study of the obtained regions that related the axial dispersion with the other parameters, as illustrated in Figure 15, reveals that the objective function values decrease as the value of the D_{ax} parameter also decreases. Consequently, it was inferred that the pre-defined boundaries for this parameter were larger than necessary, as higher values of axial dispersion do fail to yield satisfactory objective function values. To optimize subsequent estimations, the upper boundary for axial dispersion was reduced by an order of magnitude.

Lastly, revisions were made to the boundaries of the size exclusion constant for BSA, $K_{SEC,BSA}$. From the particle regions, it was inferred that the upper boundary exceeded the necessary range since objective function values diminished as the parameter value approached zero. This occurrence can be seen in some examples presented in Figure 16. Consequently, to simplify the optimization process and decrease the computational demands, the upper limit for this parameter was decreased in subsequent estimations.

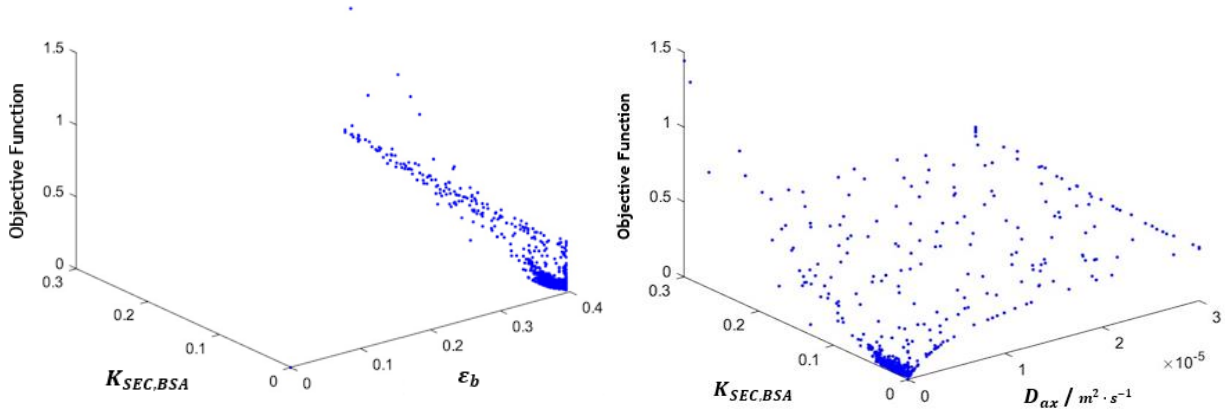

Figure 16 - Particle density regions for the parameter $K_{SEC,BSA}$, with ε_b (left) and D_{ax} (right), compared to the objective function value yielded

The parameters not mentioned in these particle density region observations, ε_b and $K_{SEC,Mb}$, did not suffer any limit alterations since their pre-defined values were assumed as appropriate for further estimations.

4.2 Estimability Analysis

To disregard the mass transfer parameters from the total parameter estimation, an estimability analysis was conducted for each of the fixed bed experiments. This analysis aimed to quantify and compare the impact of these parameters in the final result with others in the parameter set. Moreover, it sought to detect significant correlations among parameters that would hinder the parameter estimation process.

The magnitude table generated during the estimability analysis shows the correlation between parameters and their influence on the estimation process. The resulting magnitude tables are presented below, Table 4 and Table 5 for the single experiments, and Table 6 for the binary experiments. Highlighted cells correspond to the columns (parameters) with the highest magnitudes for each iteration, serving as references for the orthogonalization in that particular iteration:

Table 4 - Magnitudes at each iteration for the Myoglobin fixed-bed experiments

Iteration	ε_p	D_{ax}	Q_{feed}	K_{SEC}	k_h
1	5.35×10^{7}	3.36×10^{6}	1.21×10^{9}	6.63×10^{3}	7.45
2	8.00×10^{6}	3.33×10^{6}	1.02×10^{-23}	6.62×10^{3}	7.48
3	1.20×10^{-24}	3.20×10^{6}	6.84×10^{-24}	5.96×10^{3}	7.06
4	1.19×10^{-24}	1.92×10^{-24}	6.82×10^{-24}	5.96×10^{3}	6.24
5	1.18×10^{-24}	1.91×10^{-24}	6.77×10^{-24}	3.63×10^{-28}	6.04

Table 5 - Magnitudes at each iteration for the BSA fixed-bed experiments

Iteration	ε_p	D_{ax}	Q_{feed}	K_{SEC}	k_h
1	8.34×10^{6}	3.82×10^{6}	1.19×10^{7}	2.40×10^{3}	0.416
2	5.68×10^{4}	1.58×10^{6}	1.91×10^{-24}	2.38×10^{3}	0.138
3	5.38×10^{4}	1.70×10^{-25}	1.91×10^{-24}	2.27×10^{3}	0.132
4	1.67×10^{-27}	1.68×10^{-25}	1.91×10^{-24}	7.32×10^{2}	0.018
5	1.56×10^{-27}	1.65×10^{-25}	1.90×10^{-24}	1.01×10^{-28}	0.017

Table 6 - Magnitudes at each iteration for the binary fixed-bed experiments

Iteration	ε_p	D_{ax}	Q_{feed}	$K_{SEC,BSA}$	$K_{SEC,Mb}$	$k_{h,BSA}$	$k_{h,Mb}$
1	2.97×10^{6}	9.51×10^{7}	2.66×10^{6}	2.55×10^{6}	4.96×10^{4}	0.919	23.6
2	2.97×10^{6}	3.61×10^{-24}	2.63×10^{6}	2.51×10^{6}	4.94×10^{4}	0.904	23.5
3	1.63×10^{-27}	3.61×10^{-24}	8.42×10^{5}	1.67×10^{6}	4.71×10^{4}	0.671	22.4
4	1.61×10^{-27}	3.60×10^{-24}	6.59×10^{5}	1.50×10^{-25}	4.50×10^{4}	0.436	22.1
5	1.60×10^{-27}	3.57×10^{-24}	2.60×10^{-26}	1.49×10^{-25}	4.21×10^{4}	0.389	22.1
6	1.60×10^{-27}	3.57×10^{-24}	2.57×10^{-26}	1.49×10^{-25}	4.39×10^{-27}	0.389	17.6

Across all experiments, it is possible to observe a modest correlation between parameters, particularly between the bed porosity, the size exclusion constants, and the feed flow. This correlation is evident from the slight variation in the magnitude of the columns associated with

the correlated parameters whenever the sensitivity matrix, S, is orthogonalized. Nevertheless, this variation was deemed insignificant, so no correlation that would inhibit a successful and correct estimation was identified. As mentioned in the previous chapter, the cut-off for all the experiments had a defined value of 100. Consequently, any magnitude below the cut-off value was not considered as the reference for orthogonalization and would imply the end of the estimability analysis.

As it was expected, the magnitude of the column corresponding to the mass transfer coefficient, k_h, for all experiments, consistently fell below the established cut-off value in each iteration. This finding indicates that the influence of this parameter in the final result, *i.e.*, in the line formed by the mathematical model, is nearly negligible, corroborating the speculation postulated before.

To ensure that removing these parameters from subsequent estimations would still yield a global minimum for the objective function, an analysis involving the Hessian matrix was performed. Following the construction of the sensitivity matrices, S_f, excluding the k_h parameters, the corresponding Hessian matrices were generated. Subsequently, the eigenvalues of these Hessian matrices were computed. Table 7 displays the eigenvalues generated from each experiment.

Table 7 - Eigenvalues of the Hessian matrices obtained for each fixed-bed experiment

Mb Experiment	BSA Experiment	Binary Experiment
1.19×10^4	1.40×10^3	6.39×10^4
6.26×10^6	6.49×10^4	6.18×10^6
1.70×10^7	2.88×10^6	2.33×10^7
2.52×10^9	4.51×10^7	5.45×10^7
		1.54×10^8

The positive nature of all generated eigenvalues indicates that the Hessian matrices are positive-definite and provide mathematical proof that a global minimum exists within the region defined by the estimable parameters.

Based on the examination of the results obtained thus far, it was determined that the mass transfer coefficients, k_h, would be assigned fixed values obtained from the fixed bed parameter estimation and were removed from estimation in the total parameter estimation process. So, $k_{h,BSA}$ will have a fixed value of 0.566 s^{-1}, while $k_{h,Mb}$ will have a fixed value of 0.047 s^{-1}.

4.3 Total Parameter Estimation

Upon implementing all necessary changes, the total parameter estimation was conducted. This study consisted of the estimation of parameters for the SMB experiment while utilizing the binary mixture fixed bed experiments to aid the estimation. The primary goal of this "global" method is, if possible, to identify a set of parameters that could generate a mathematical model that is capable of simultaneously fitting both the SMB experimental data and the fixed bed experimental data. Additionally, a Blue Dextran fixed bed experiment was also incorporated into this estimation to enhance the estimation of the fixed bed porosity ε_b and, consequently, to improve the estimation of the size exclusion constants, K_{SEC}, due to the observed correlation between these parameters. A Particle Swarm Optimization algorithm using 50 particles and 30 iterations was applied.

Table 8 presents the parameters obtained from Rios' calculations, the estimated parameters and their respective objective function values. Figure 17 illustrates a graphical comparison between the SMB experimental data and the corresponding mathematical models for the Extract (left) and Raffinate (right) streams, using each set of parameters.

Table 8 - Rios' and Estimated set of parameters for the global method

	Rios' Parameters	Estimated Parameters
$\boldsymbol{\varepsilon_b}$	0.39	0.394
$\boldsymbol{D_{ax}\ (m^2 \cdot s^{-1})}$	3.83 x 10^{-7}	9.89 x 10^{-8}
$\boldsymbol{Q_{feed}\ (m^3 \cdot s^{-1})}$	3.17 x 10^{-8}	3.21 x 10^{-8}
$\boldsymbol{K_{SEC,BSA}}$	0	0.029
$\boldsymbol{K_{SEC,Mb}}$	0.3	0.302
$\boldsymbol{Q_F\ (l \cdot\ min^{-1})}$	0.85 x 10^{-3}	0.865 x 10^{-3}
$\boldsymbol{Q_D\ (l \cdot\ min^{-1})}$	5.000 x 10^{-3}	5.005 x 10^{-3}
$\boldsymbol{Q_E\ (l \cdot\ min^{-1})}$	2.75 x 10^{-3}	2.768 x 10^{-3}
$\boldsymbol{Q_{IV}\ (l \cdot\ min^{-1})}$	1.5 x 10^{-3}	1.509 x 10^{-3}
$\boldsymbol{f_{obj}}$	0.0055	0.0014

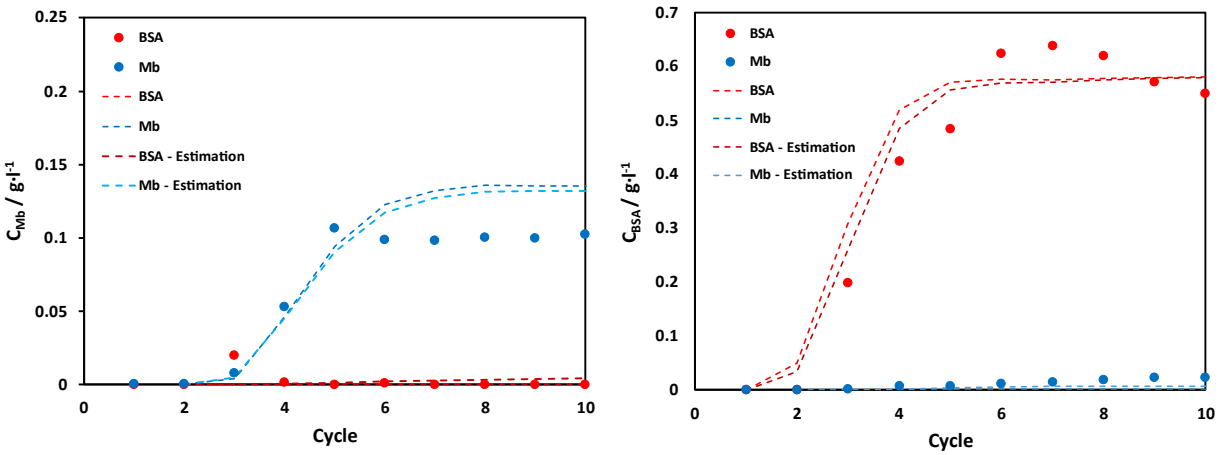

Figure 17 - Graphical comparison between SMB experimental data (points) and the models with Rios' parameters and estimated parameters at the extract (left) and raffinate (right)

Upon a first analysis of the graphical comparisons, it may be difficult to ascertain which model yields a more significant adjustment to the experimental points. This observation is especially complicated since some experimental points are better adjusted by the model with estimated parameters, while others seem to be better aligned with Rios' parameters model. However, it is important to note that the concentration range, represented along the y-axis, varies significantly between the graphs; therefore, initial interpretations can be misleading. The objective functions compiled in Table 8 suggest that the estimated model has a better alignment to the experimental data when comparing it to the model with Rios' parameters, due to the decrease in objective function values. By scaling the y-axis divisions in the right graph to ensure consistent representations in both the extract and raffinate streams, the decrease in objective function value becomes more justified. This can be observed in Figure 18, where it is clearer that the estimated model aligns more accurately with the experimental points:

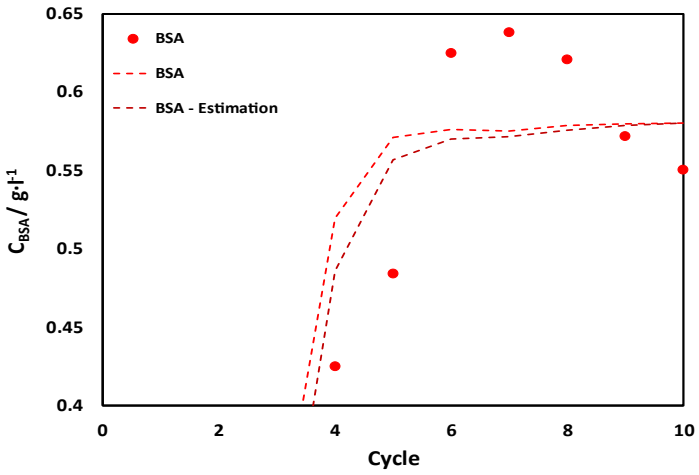

Figure 18 – Zoom in of the graphical comparison between SMB experimental data (points) and the models with Rios' parameters and estimated parameters at the raffinate port

Therefore, the total estimation provided a set of parameters that created a model which exhibited a slightly better adjustment to the experimental points at the extract and raffinate ports than the model generated with Rios' parameters. Hence, the overall residuals obtained with the estimated parameters were lower.

The main difference between parameter sets was observed, once again, in the axial dispersion coefficient, D_{ax}, which reported a slight decrease in its value. The operational flow rates also revealed some slight variations that could have influenced the final result. The values for the size exclusion constants, K_{SEC} displayed some deviations, mainly the parameter relating to the BSA protein. However, these deviations were expected as previous estimations had reported similar findings. A slight increase in the fixed bed porosity was also observed, although previous experiments have indicated that this deviation does not affect the mathematical model as much as other variations might.

Other parameters were also estimated, namely the bed porosity and the Peclet number of each column, $\varepsilon_{b,n}$ and Pe_n, respectively, in which n is the column number. However, the interest in the estimation of these parameters was the study of their uncertainty. As an informative compilation, in Table 9 below, the obtained values for these parameters are compared with the original ones proposed by Rios:

Table 9 – Continuation of the estimated parameters in the total estimation

	Rios' Parameters	Estimated Parameters		Rios' Parameters	Estimated Parameters
$\boldsymbol{\varepsilon_{b,1}}$	0.390	0.389	$\boldsymbol{Pe_1}$	67	50.00
$\boldsymbol{\varepsilon_{b,2}}$	0.390	0.389	$\boldsymbol{Pe_2}$	67	62.38
$\boldsymbol{\varepsilon_{b,3}}$	0.390	0.390	$\boldsymbol{Pe_3}$	67	50.02
$\boldsymbol{\varepsilon_{b,4}}$	0.390	0.390	$\boldsymbol{Pe_4}$	67	96.55
$\boldsymbol{\varepsilon_{b,5}}$	0.390	0.389	$\boldsymbol{Pe_5}$	67	85.36
$\boldsymbol{\varepsilon_{b,6}}$	0.390	0.389	$\boldsymbol{Pe_6}$	67	50.12

An absolute and relative comparison was performed between these sets of parameters to validate the obtained results and assess the differences between the values obtained in the total parameter study and the reference parameters from the fixed bed studies. For the myoglobin fixed-bed experiment, the deviations are presented in Table 10, and a graphical representation illustrating these differences is shown in Figure 19.

Table 10 - Compilation of deviation between individual and global studies for the Mb experiment

Parameter	Absolute Deviation	Relative Deviation (%)
ε_b	1.44 x 10^{-3}	0.37
D_{ax} $(m^2 \cdot s^{-1})$	2.29 x 10^{-7}	69.8
Q_{feed} $(m^3 \cdot s^{-1})$	3.21 x 10^{-10}	0.99
K_{SEC}	3.09 x 10^{-2}	9.30

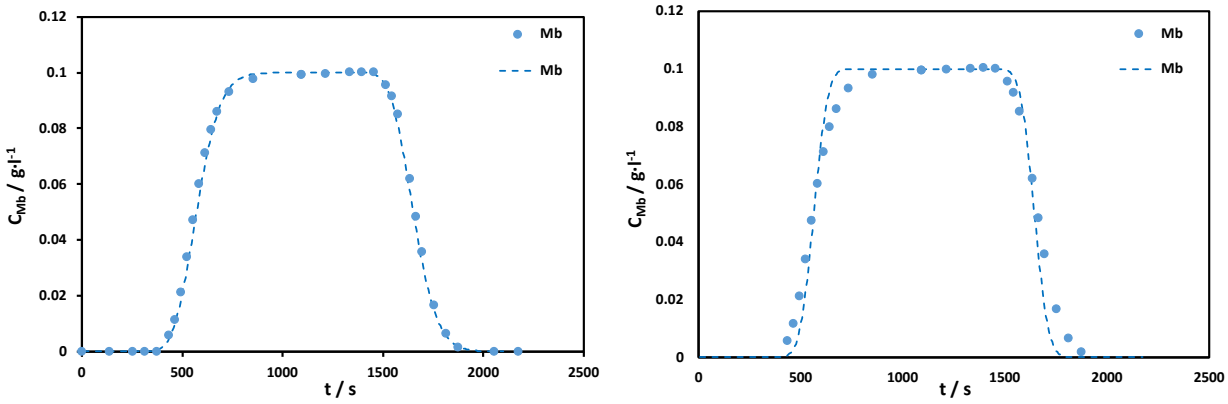

Figure 19 - Graphical comparison between fixed-bed reference parameters (left) and estimated parameters yielded by the global study (right) – Mb experiment

The global study resulted in a significant difference in axial dispersion, resulting in an approximate 70 % decrease compared to the reference value. However, it is important to note that the estimation range for this parameter is considerable (spanning three orders of magnitude), and the reference value itself is very small; therefore, a 70 % deviation does not change the mathematical model as much as the variation might induce. Nevertheless, differences between the obtained mathematical models can be observed in the graphical representation, particularly in the time intervals of 600 to 900 seconds and 1600 to 1900 seconds, where the effect of this axial dispersion deviation is evident. The other parameters obtained through the global parameter estimation show close similarity to the reference values. Lastly, the objective function values of all the myoglobin experiments were compared to validate further the results obtained. The respective values are presented in Table 11 below:

Table 11 - Objective functions between Mb experiments

	Rios' Parameters	Fixed Bed Experiment	Total Experiment
f_{obj}	1.23 x 10^{-5}	5.01 x 10^{-6}	5.66 x 10^{-5}

Consistent with the visual analysis, it is evident that the parameters obtained in the total parameter study resulted in a model that exhibited a worse alignment to the experimental data than the original mathematical model with Rios' parameters, as indicated by the higher objective function value. Nonetheless, due to the low value of the objective function, the total estimation was considered to have a reasonable adjustment to the experimental data.

Subsequently, a similar investigation was conducted for the BSA protein. Table 12 provides an overview of the deviations, while Figure 20 illustrates the graphical differences between reference and "globally" estimated parameters:

Table 12 - Compilation of deviation between individual and global studies for the BSA experiment

Parameter	Absolute Deviation	Relative Deviation (%)
ε_b	2.22 x 10^{-3}	0.57
D_{ax} $(m^2 \cdot s^{-1})$	4.49 x 10^{-8}	83.1
Q_{feed} $(m^3 \cdot s^{-1})$	4.93 x 10^{-10}	1.56
K_{SEC}	2.19 x 10^{-4}	0.99

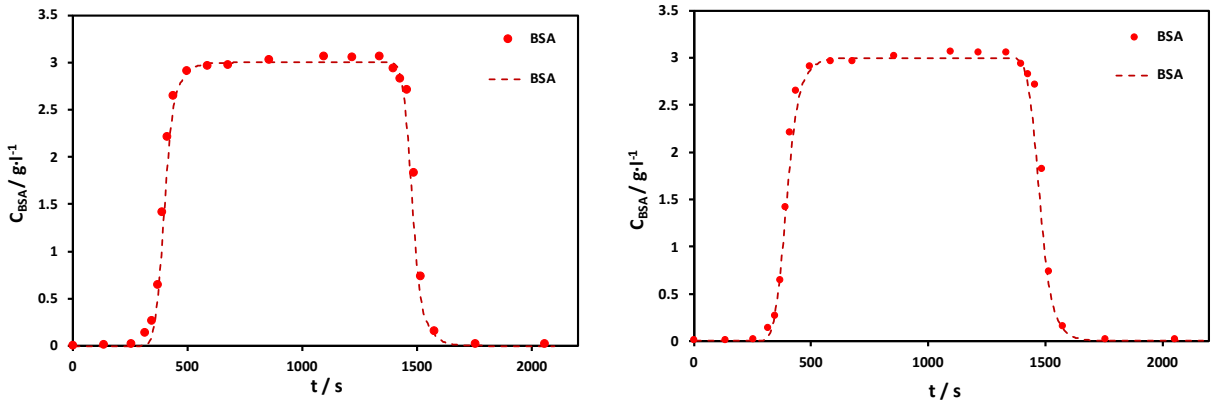

Figure 20 - Graphical comparison between fixed-bed reference parameters (left) and estimated parameters yielded by the global study (right – BSA experiment

Upon examining Table 12, it is possible to infer that axial dispersion deviates significantly from the reference parameter, reporting a deviation of about 83 %. But, since the value of the reference parameter is very small, this difference will be negligible in the resulting mathematical model. This observation is further supported by the graphical representations, which exhibit no noticeable divergences. In conjunction with the insignificant deviations shown by other parameters, these findings suggest that the total parameter estimation, in this case,

provides an accurate adjustment to the experimental data. By, once again, comparing objective function values present in Table 13, it is possible to substantiate this claim:

Table 13 - Objective functions between BSA experiments

	Rios' Parameters	Individual Experiment	Total Experiment
f_{obj}	0.146	0.0326	0.0414

In contrast to the myoglobin experiment, the model with Rios' parameters performs the worst. Unsurprisingly, the fixed bed experiment and the reference parameters generate the most accurate mathematical model for the experimental data adjustment.

Lastly, a similar comparison and analysis for the binary experiment was conducted.

Table 14 - Compilation of deviation between individual and global studies for the binary experiment

Parameter	Absolute Deviation	Relative Deviation (%)
ε_b	1.40 x 10^{-3}	0.36
D_{ax} $(m^2 \cdot s^{-1})$	1.43 x 10^{-9}	1.47
Q_{feed} $(m^3 \cdot s^{-1})$	4.79 x 10^{-10}	1.51
$K_{SEC,BSA}$	1.31 x 10^{-2}	84.4
$K_{SEC,Mb}$	4.15 x 10^{-3}	1.39

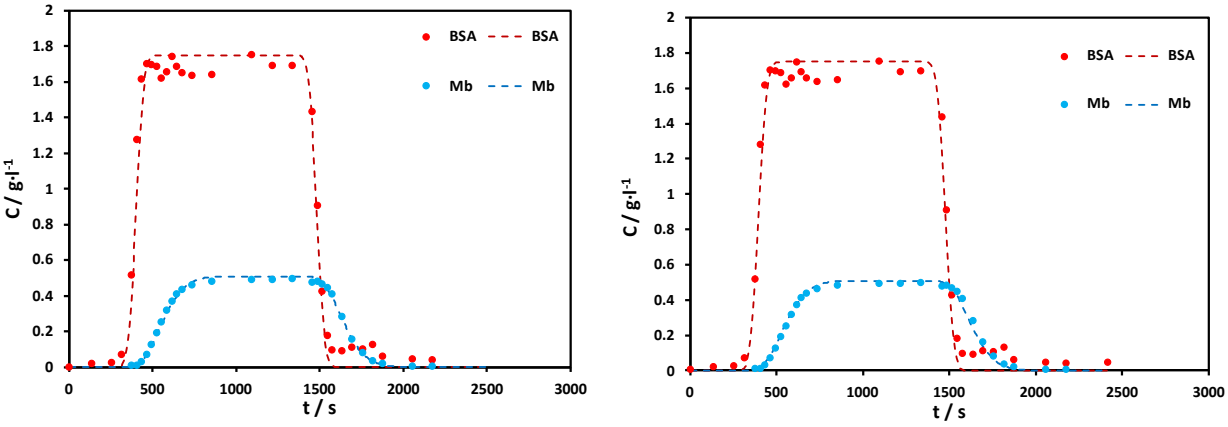

Figure 21 - Graphical comparison between fixed-bed reference parameters (left) and estimated parameters yielded by the global study (right) – binary Exp. 1

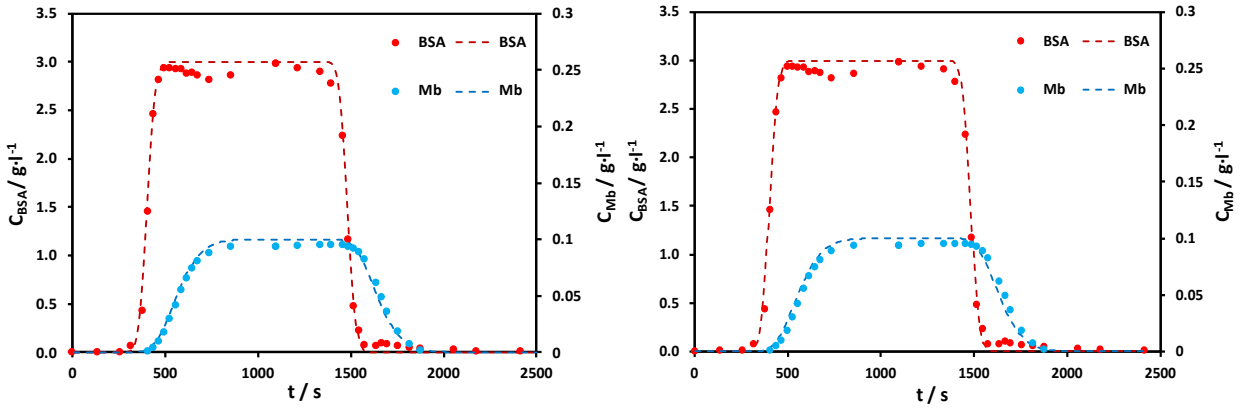

Figure 22 - Graphical comparison between fixed-bed reference parameters (left) and estimated parameters yielded by the global study (right) – binary Exp. 2

In the binary experiment, the axial dispersion did not suffer a deviation akin to the one observed in the experiments seen previously. The parameter that showed a notable variation was the size exclusion constant for BSA, $K_{SEC,BSA}$, but considering the small absolute increase of 1.31 x 10^{-2} and the nature of this parameter, this minor variation will not have a significant impact on the alignment of the mathematical model. The remaining parameters also displayed minor variations compared to the reference values. Hence, it is possible to conclude that the model obtained using the total estimated parameters is well-adjusted to the experimental data. To further validate this claim, a comparison between objective function values between all the binary experiments has been, once again, performed. The values are presented in Table 15:

Table 15 - Objective functions between binary experiments

	Rios' Parameters	Individual Experiment	Total Experiment
f_{obj}	0.0151	0.0107	0.0112

Once again, the mathematical model with Rios' parameters performed the poorest among all the models. This reaffirms the hypothesis that the parameters estimated with the total array of experiments align well with the experimental data in all the fixed bed experiments of the case study, except for the myoglobin experiment, where a slight deviation can be observed.

The obtained set of total parameters was considered satisfactory since these parameters generate a model that has an adequate alignment to the myoglobin fixed bed and SMB experimental data, and an exceptional adjustment to the BSA and binary fixed bed experimental data, simultaneously. Therefore, no further estimations were deemed necessary to improve the quality of the parameter set.

4.4 Uncertainty Analysis

The uncertainty of the parameters that were estimated soon followed the total estimation. Using expression 3.23, it was possible to compute the uncertainties for each parameter. They are compiled in Table 16, alongside the mean values of each parameter, obtained through the Particle Swarm Optimization:

Table 16 - Parameter Uncertainties

	Mean value	Uncertainty		Mean value	Uncertainty
ε_b	0.3934	± 0.0022	Pe_6	53.917	± 18.395
D_{ax} $(m^2 \cdot s^{-1})$	1.773 x 10^{-7}	± 7.368 x 10^{-7}	$\varepsilon_{b,1}$	0.3894	± 0.0005
Q_f $(m^3 \cdot s^{-1})$	3.197 x 10^{-8}	± 5.407 x 10^{-10}	$\varepsilon_{b,2}$	0.3893	± 0.0005
$K_{SEC,BSA}$	0.0267	± 0.0316	$\varepsilon_{b,3}$	0.3902	± 0.0008
$K_{SEC,Mb}$	0.3104	± 0.0931	$\varepsilon_{b,4}$	0.3900	± 0.0010
Pe_1	50.800	± 9.069	$\varepsilon_{b,5}$	0.3892	± 0.0005
Pe_2	62.523	± 9.516	$\varepsilon_{b,6}$	0.3894	± 0.0014
Pe_3	51.068	± 10.852	Q_F $(l \cdot min^{-1})$	8.647 x 10^{-4}	± 1.077 x 10^{-5}
Pe_4	94.937	± 12.964	Q_D $(l \cdot min^{-1})$	5.007 x 10^{-3}	± 3.344 x 10^{-5}
Pe_5	84.836	± 7.908	Q_E $(l \cdot min^{-1})$	2.765 x 10^{-3}	± 1.364 x 10^{-5}
			Q_{IV} $(l \cdot min^{-1})$	1.508 x 10^{-3}	± 6.390 x 10^{-6}

Upon examination of Table 16, it is evident that all the best values obtained for the parameters through the PSO fall within their respective uncertainty values, as expected. Notably, the uncertainty interval for the D_{ax} and $K_{SEC,BSA}$ parameters may encompass negative values. However, since negative axial dispersion and a negative size exclusion constant are physically infeasible, the lower limit of zero was imposed on these intervals, neglecting the negative values within these uncertainties.

As an informative study, confidence regions were built based on the Fisher-Snedecor test. As these confidence regions are not a primary focus of this work, they are included in Appendix B.

Lastly, the uncertainty propagation was computed using expression 3.25. The value obtained for the propagated uncertainty, U_c, was 0.0737. This value quantifies the overall uncertainty imposed on the final result by the estimated parameters.

5 Conclusion

The objectives of this dissertation were to evaluate the assumed mathematical model for the separation at hand, to estimate a set of parameters that would create a model that adjusts to the experimental data available, and to perform estimability and uncertainty evaluations to complete, validate and justify the results obtained.

The adsorption model assumed by Rios for the Sephadex G-50 size exclusion separation, a linear model, was considered ideal since size exclusion has a very simple mechanism that resembles adsorption, where no competition is found, and the affinity of the proteins to the material only depends on the size of the protein.

The parameter estimation for the individual fixed-bed experiments, using 100 iterations and 100 particles, yielded a set of reference parameters for each experiment, with close to an optimal alignment to the experimental data.

The following estimability and particle density regions analysis determined that the parameter mass transfer coefficient, k_h, for both myoglobin and BSA, does not influence the final result in the chosen parameter boundaries. Therefore, its estimation was considered irrelevant and was removed from further studies. Observing different particle density regions, some parameters' upper and lower limits were modified, so that an accurate estimation, with less computational power required, could be performed.

The parameter estimation using all the experiments available, SMB experiments included, and using 30 iterations and 50 particles, resulting in a set of parameters that created a model with satisfactory adjustment to the SMB experimental data. Using the reference values obtained during the individual fixed-bed studies, it was possible to observe slight differences in some parameters, namely the axial dispersion coefficient, D_{ax}, for the myoglobin and BSA individual experiments and the size exclusion constant for BSA, $K_{SEC,BSA}$, in the binary fixed-bed experiment. Nonetheless, a satisfactory adjustment was revealed for the myoglobin experiment and a close-to-optimal adjustment for the BSA and binary experiments.

The uncertainty of each of the estimated parameters, and the propagated uncertainty in the final result, was computed to obtain confidence intervals and characterize the results obtained.

The methodology considered and used in this work proved to be a successful course of action to estimate parameters in fixed bed and SMB environments and revealed to be a reliable technique to validate the results obtained; therefore, this approach may have the capability of aiding the research studies performed in the realm of parameter estimation.

6 Assessment of the work done

6.1 Objectives Achieved

The evaluation of the assumed mathematical adsorption models confirmed their appropriateness for the present separation. The objective of using Particle Swarm Optimization to generate a set of parameters that aligns the resulting mathematical model to the experimental data for various experiments was successful. Furthermore, the estimability and uncertainty analysis were conducted as planned and completed satisfactorily.

6.2 Other Work Carried Out

Alongside the work conducted for this dissertation, similar studies to those described in this document were performed using other solid phase materials, namely Q Sepharose FF and Hydroxyapatite, for the same myoglobin and BSA separation. Albertina Rios also used these adsorbents during her Ph.D. thesis.

Regarding the material Q Sepharose FF, an ionic exchange matrix, the gPROMS models that describe the adsorption, as well as the SMB equipment, were adapted for parameter estimation. In addition, several MATLAB® scripts were built to prepare the parameter estimation for the fixed bed experiments.

In the case of the Hydroxyapatite material, a ceramic matrix, a similar work to that performed for the Q Sepharose FF material was performed. Furthermore, the initial parameter estimation for the individual fixed-bed experiments was also completed; however, the results obtained did not meet the desired level of satisfaction.

6.3 Final Assessment

Several improvements can be made to the developed work. The increase in the number of iterations and particles used in the PSO could enhance the parameters quality, the particle density region, and uncertainty analysis due to the higher number of particles to observe. A thorough uncertainty analysis could provide valuable insights not obtained in this dissertation.

Following what was affirmed in the last sub-section, a suggestion for further work is the replication of the same study conducted in this dissertation for the remaining two materials, more complex in nature. The use of other optimization methods could also be an interesting investigation. By comparing computational time and power required to reach similar results, it would be possible to identify the most efficient and effective method for estimation.

7 References

1. Nehete, J.; Bhambar, R.; Narkhede, M.; Gawali, S. J. P. r., Natural proteins: Sources, isolation, characterization and applications. **2013,** *7* (14), 107.
2. Smith, D. M., Protein Separation and Characterization Procedures. In *Food Analysis*, Nielsen, S. S., Ed. Springer International Publishing: Cham, 2017; pp 431-453.
3. Sing, K. S. J. P.; chemistry, a., Reporting physisorption data for gas/solid systems with special reference to the determination of surface area and porosity (Recommendations 1984). **1985,** *57* (4), 603-619.
4. Rios, A. Protein Separation and Purification by Cyclic Adsorption Processes. 2023.
5. Gomes, P. S.; Zabkova, M.; Zabka, M.; Minceva, M.; Rodrigues, A. E., Separation of chiral mixtures in real SMB units: The FlexSMB-LSRE®. **2010,** *56* (1), 125-142.
6. Whitford, D., *Proteins: structure and function*. John Wiley & Sons: 2013.
7. Lodish, H.; Berk, A.; Kaiser, C. A.; Kaiser, C.; Krieger, M.; Scott, M. P.; Bretscher, A.; Ploegh, H.; Matsudaira, P., *Molecular cell biology*. Macmillan: 2008.
8. Kumar, M.; Tomar, M.; Punia, S.; Dhakane-Lad, J.; Dhumal, S.; Changan, S.; Senapathy, M.; Berwal, M. K.; Sampathrajan, V.; Sayed, A. A. S.; Chandran, D.; Pandiselvam, R.; Rais, N.; Mahato, D. K.; Udikeri, S. S.; Satankar, V.; Anitha, T.; Reetu; Radha; Singh, S.; Amarowicz, R.; Kennedy, J. F., Plant-based proteins and their multifaceted industrial applications. *LWT* **2022,** *154*, 112620.
9. Hey, T.; Fiedler, E.; Rudolph, R.; Fiedler, M., Artificial, non-antibody binding proteins for pharmaceutical and industrial applications. *Trends in Biotechnology* **2005,** *23* (10), 514-522.
10. Ryu, H.; Fuwad, A.; Yoon, S.; Jang, H.; Lee, J. C.; Kim, S. M.; Jeon, T.-J., Biomimetic Membranes with Transmembrane Proteins: State-of-the-Art in Transmembrane Protein Applications. **2019,** *20* (6), 1437.
11. Nfor, B. K.; Ahamed, T.; van Dedem, G. W.; van der Wielen, L. A.; van de Sandt, E. J.; Eppink, M. H.; Ottens, M., Design strategies for integrated protein purification processes: challenges, progress and outlook. **2008,** *83* (2), 124-132.
12. Saxena, A.; Tripathi, B. P.; Kumar, M.; Shahi, V. K., Membrane-based techniques for the separation and purification of proteins: An overview. *Advances in Colloid and Interface Science* **2009,** *145* (1), 1-22.
13. Bonner, P.; Bonner, P. L., *Protein purification*. Taylor & Francis: 2007.
14. Liu, S.; Li, Z.; Yu, B.; Wang, S.; Shen, Y.; Cong, H., Recent advances on protein separation and purification methods. *Advances in Colloid and Interface Science* **2020,** *284*, 102254.
15. Yehl, C. J.; Zydney, A. L., High Performance Countercurrent Membrane Purification for protein separations. *Journal of Membrane Science* **2021,** *633*, 119396.

16. Chen, S.-T.; Wickramasinghe, S. R.; Qian, X., Electrospun Weak Anion-Exchange Fibrous Membranes for Protein Purification. **2020,** *10* (3), 39.

17. Arefi-Oskoui, S.; Khataee, A.; Jabbarvand Behrouz, S.; Vatanpour, V.; Haddadi Gharamaleki, S.; Orooji, Y.; Safarpour, M., Development of MoS2/O-MWCNTs/PES blended membrane for efficient removal of dyes, antibiotic, and protein. *Separation and Purification Technology* **2022,** *280*, 119822.

18. Jungbauer, A.; Hahn, R., Chapter 22 Ion-Exchange Chromatography. In *Methods in Enzymology*, Burgess, R. R.; Deutscher, M. P., Eds. Academic Press: 2009; Vol. 463, pp 349-371.

19. Asenjo, J. A.; Andrews, B. A., Protein purification using chromatography: selection of type, modelling and optimization of operating conditions. **2009,** *22* (2), 65-76.

20. Ettre, L. S. J. L. G. N. A., MS Tswett and the invention of chromatography. **2003,** *21* (5), 458-467.

21. Lundanes, E.; Reubsaet, L.; Greibrokk, T., *Chromatography: basic principles, sample preparations and related methods*. John Wiley & Sons: 2013.

22. Coskun, O. J. N. c. o. I., Separation techniques: chromatography. **2016,** *3* (2), 156.

23. Taylor, L. T., Supercritical fluid chromatography for the 21st century. *The Journal of Supercritical Fluids* **2009,** *47* (3), 566-573.

24. Xu, Z.; Cai, J.-g.; Pan, B.-c., Mathematically modeling fixed-bed adsorption in aqueous systems. *Journal of Zhejiang University SCIENCE A* **2013,** *14* (3), 155-176.

25. Wang, J.; Guo, X., Adsorption isotherm models: Classification, physical meaning, application and solving method. *Chemosphere* **2020,** *258*, 127279.

26. Abin-Bazaine, A.; Trujillo, A. C.; Olmos-Marquez, M., Adsorption Isotherms: Enlightenment of the Phenomenon of Adsorption. In *Wastewater Treatment*, IntechOpen: 2022.

27. Forssén, P.; Fornstedt, T., A model free method for estimation of complicated adsorption isotherms in liquid chromatography. *Journal of Chromatography A* **2015,** *1409*, 108-115.

28. Broughton, D. B.; Gerhold, C. G., Continuous sorption process employing fixed bed of sorbent and moving inlets and outlets. Google Patents: 1961.

29. Juza, M.; Mazzotti, M.; Morbidelli, M., Simulated moving-bed chromatography and its application to chirotechnology. *Trends in Biotechnology* **2000,** *18* (3), 108-118.

30. Charton, F.; Nicoud, R.-M., Complete design of a simulated moving bed. *Journal of Chromatography A* **1995,** *702* (1), 97-112.

31. Faria, R. P.; Gonçalves, J. C.; Rodrigues, A. E., Advances in Simulated Moving Bed Technology. In *Advances in Chromatography Volume 57*, CRC Press: 2020; pp 163-236.

32. Sá Gomes, P.; Rodrigues, A. E., Simulated Moving Bed Chromatography: From Concept to Proof-of-Concept. **2012,** *35* (1), 17-34.

33. Faria, R. P. V.; Rodrigues, A. E., Instrumental aspects of Simulated Moving Bed chromatography. *Journal of Chromatography A* **2015,** *1421*, 82-102.

34. Pinto, J. C.; Schwaab, M., *Análise de Dados Experimentais: I. Fundamentos de Estatística e Estimação de Parâmetros*. Editora E-papers: 2007.

35. Lin, M.-H.; Tsai, J.-F.; Yu, C.-S., A Review of Deterministic Optimization Methods in Engineering and Management. *Mathematical Problems in Engineering* **2012,** *2012*, 756023.

36. Martí, R.; Reinelt, G., Heuristic Methods. In *Exact and Heuristic Methods in Combinatorial Optimization: A Study on the Linear Ordering and the Maximum Diversity Problem*, Martí, R.; Reinelt, G., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2022; pp 27-57.

37. Rao, S. S., *Engineering optimization: theory and practice*. John Wiley & Sons: 2019.

38. Coelho, L. d. S.; Sierakowski, C. A., A software tool for teaching of particle swarm optimization fundamentals. *Advances in Engineering Software* **2008,** *39* (11), 877-887.

39. Kennedy, J.; Eberhart, R. In *Particle swarm optimization*, Proceedings of ICNN'95 - International Conference on Neural Networks, 27 Nov.-1 Dec. 1995; 1995; pp 1942-1948 vol.4.

40. Venter, G.; Sobieszczanski-Sobieski, J., Particle Swarm Optimization. **2003,** *41* (8), 1583-1589.

41. Shi, Y.; Eberhart, R. In *A modified particle swarm optimizer*, 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360), 4-9 May 1998; 1998; pp 69-73.

42. Xiaohui, H.; Eberhart, R. C.; Yuhui, S. In *Engineering optimization with particle swarm*, Proceedings of the 2003 IEEE Swarm Intelligence Symposium. SIS'03 (Cat. No.03EX706), 26-26 April 2003; 2003; pp 53-57.

43. Wang, D.; Tan, D.; Liu, L., Particle swarm optimization algorithm: an overview. *Soft Computing* **2018,** *22* (2), 387-408.

44. Nogueira, I. B. R.; Ribeiro, A. M.; Rodrigues, A. E.; Loureiro, J. M., Dynamics of a True Moving Bed separation process: Effect of operating variables on performance indicators using orthogonalization method. *Computers & Chemical Engineering* **2016,** *86*, 5-17.

45. Nogueira, I. B. R.; Pontes, K. V., Parameter estimation with estimability analysis applied to an industrial scale polymerization process. *Computers & Chemical Engineering* **2017,** *96*, 75-86.

46. Yao, K. Z.; Shaw, B. M.; Kou, B.; McAuley, K. B.; Bacon, D. W., Modeling Ethylene/Butene Copolymerization with Multi-site Catalysts: Parameter Estimability and Experimental Design. *Polymer Reaction Engineering* **2003,** *11* (3), 563-588.

47. Kravaris, C.; Hahn, J.; Chu, Y. J. C.; engineering, c., Advances and selected recent developments in state and parameter estimation. **2013,** *51*, 111-123.

48. Lund, B. F.; Foss, B. A., Parameter ranking by orthogonalization—Applied to nonlinear mechanistic models. *Automatica* **2008,** *44* (1), 278-281.

49. Alhaj-Dibo, M.; Maquin, D.; Ragot, J., Data reconciliation: A robust approach using a contaminated distribution. *Control Engineering Practice* **2008,** *16* (2), 159-170.

50. Crowe, C. M., Data reconciliation — Progress and challenges. *Journal of Process Control* **1996,** *6* (2), 89-98.

51. Iso, I.; OIML, B. J. G., Switzerland, Guide to the Expression of Uncertainty in Measurement. **1995,** *122*, 16-17.

52. Geneva, I. O. f. S., *Guide to the Expression of uncertainty in measurement*. BIPM: 2008.

53. Suzuki, K.; Harada, H.; Sato, K.; Okada, K.; Tsuruta, M.; Yajima, T.; Kawajiri, Y., Utilization of operation data for parameter estimation of simulated moving bed chromatography. **2022,** *4* (1), e10103.

54. Santos, R. V. A.; Prudente, A. N.; Ribeiro, A. M.; Rodrigues, A. E.; Loureiro, J. M.; Martins, M. A. F.; Pontes, K. V.; Nogueira, I. B. R., Global Approach for Simulated Moving Bed Model Identification: Design of Experiments, Uncertainty Evaluation, and Optimization Strategy Assessment. *Industrial & Engineering Chemistry Research* **2021,** *60* (21), 7904-7916.

55. Schwaab, M.; Biscaia, J. E. C.; Monteiro, J. L.; Pinto, J. C., Nonlinear parameter estimation through particle swarm optimization. *Chemical Engineering Science* **2008,** *63* (6), 1542-1552.

56. He, Q.; Wang, L.; Liu, B., Parameter estimation for chaotic systems by particle swarm optimization. *Chaos, Solitons & Fractals* **2007,** *34* (2), 654-661.

57. Gill, M. K.; Kaheil, Y. H.; Khalil, A.; McKee, M.; Bastidas, L., Multiobjective particle swarm optimization for parameter estimation in hydrology. **2006,** *42* (7).

58. Ouyang, A.; Li, K.; Truong, T. K.; Sallam, A.; Sha, E. H. M., Hybrid particle swarm optimization for parameter estimation of Muskingum model. *Neural Computing and Applications* **2014,** *25* (7), 1785-1799.

59. Sakthivel, V. P.; Bhuvaneswari, R.; Subramanian, S., Multi-objective parameter estimation of induction motor using particle swarm optimization. *Engineering Applications of Artificial Intelligence* **2010,** *23* (3), 302-312.

60. Hendgen-Cotta, U. B.; Kelm, M.; Rassaf, T., Myoglobin functions in the heart. *Free Radical Biology and Medicine* **2014,** *73*, 252-259.

61. Essa, H.; Magner, E.; Cooney, J.; Hodnett, B. K., Influence of pH and ionic strength on the adsorption, leaching and activity of myoglobin immobilized onto ordered mesoporous silicates. *Journal of Molecular Catalysis B: Enzymatic* **2007,** *49* (1), 61-68.

62. Peters, T., Serum Albumin. In *Advances in Clinical Chemistry*, Bodansky, O.; Stewart, C. P., Eds. Elsevier: 1970; Vol. 13, pp 37-111.

63. Peters Jr, T., *All about albumin: biochemistry, genetics, and medical applications*. Academic press: 1995.

64. PDB, R. 3RGK - Crystal Structure of Human Myoglobin Mutant K45R. https://www.rcsb.org/3d-view/3RGK/1.

65. PDB, R. 4F5S - Crystal Structure of Bovine Serum Albumin. https://www.rcsb.org/3d-view/4F5S/1.

66. Rios, A. G.; Ribeiro, A. M.; Rodrigues, A. E.; Ferreira, A. F. P., Bovine serum albumin and myoglobin separation by size exclusion SMB. *Journal of Chromatography A* **2020,** *1628*, 461431.

Appendix A - Particle Density Regions

A.1. Particle Density Regions for the Myoglobin Experiment:

There were 5 parameters that were chosen to undergo estimation: ε_b, D_{ax}, Q_{feed}, K_{SEC} and k_h.

The analysis of these graphs, aided by the estimability evaluation, led to the removal of one of the parameters in further estimations, k_h. Additionally, they also were a reliable tool in the adjustment of the parameter search region limits for the total parameter estimation. Figure A.1 to Figure A.10 provide information about the particles' positions during the PSO algorithm while also presenting the objective function associated with such positions:

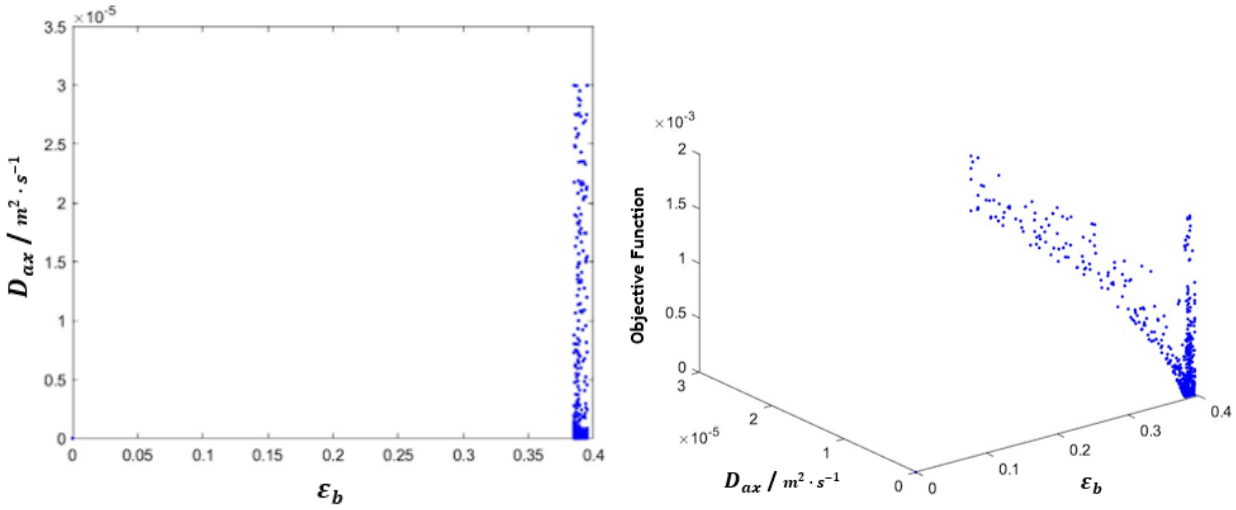

Figure A.1 - Particle density regions for the parameters ε_b and D_{ax}

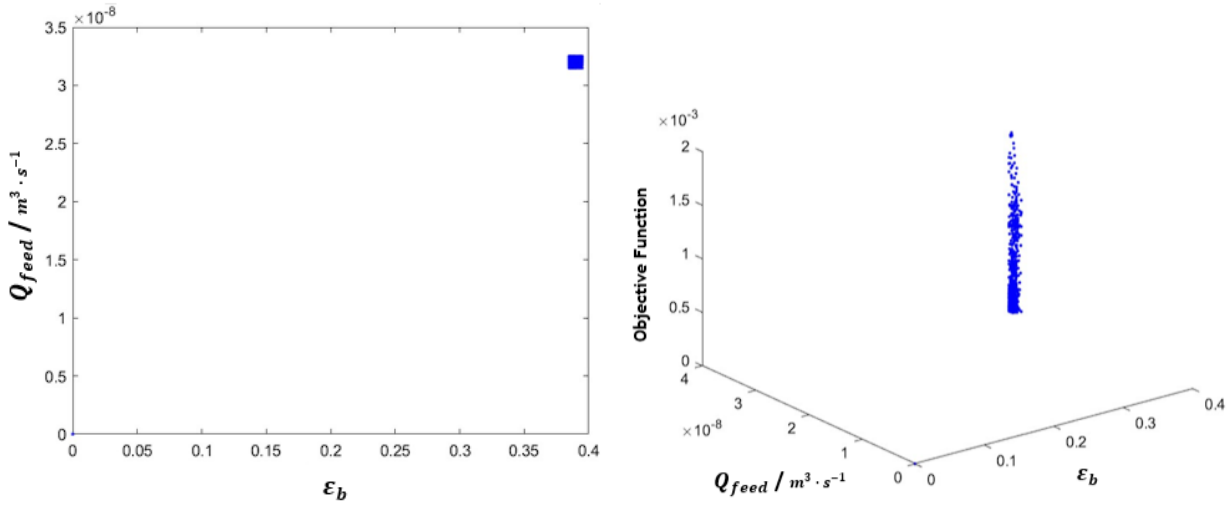

Figure A. 2 - Particle density regions for the parameters ε_b and Q_{feed}

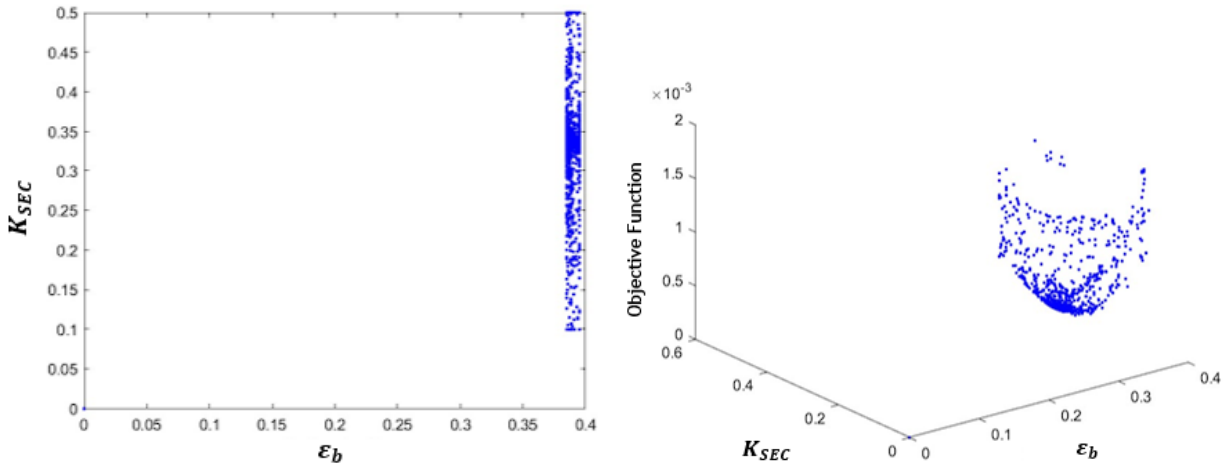

Figure A. 3 - Particle density regions for the parameters ε_b and K_{SEC}

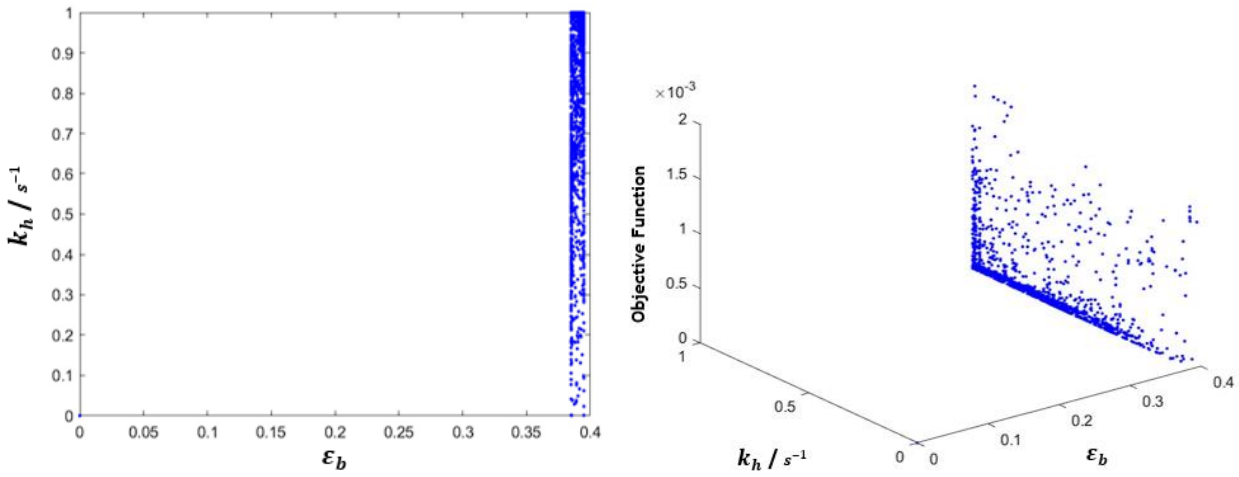

Figure A. 4 - Particle density regions for the parameters ε_b and k_h

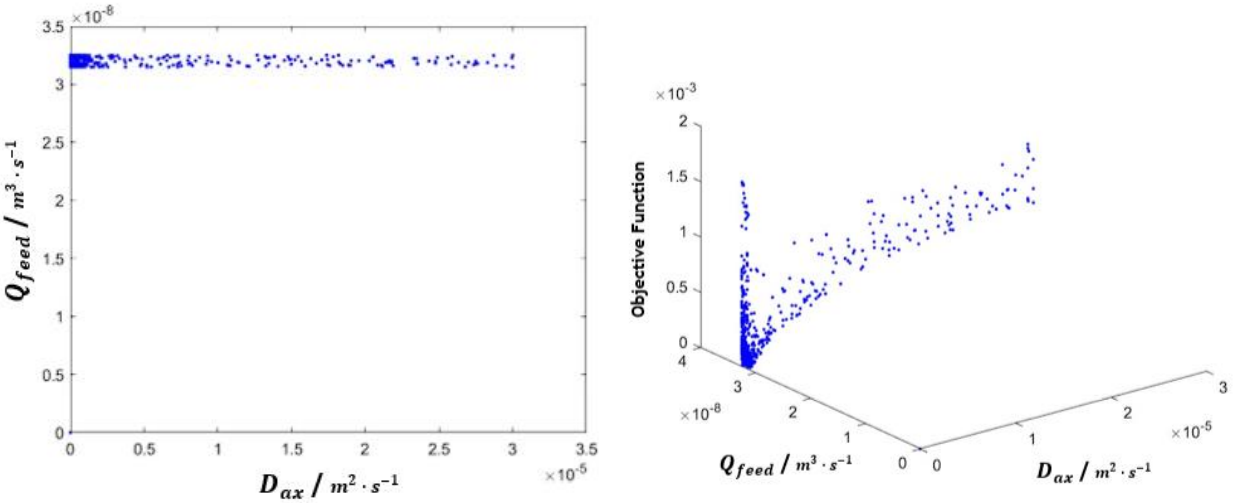

Figure A. 5 - Particle density regions for the parameters D_{ax} and Q_{feed}

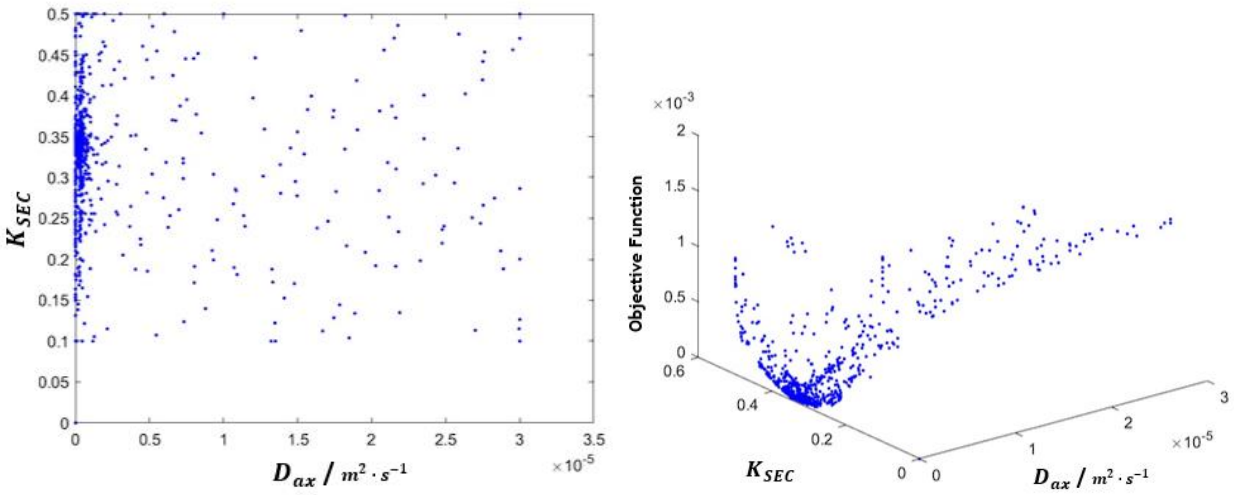

Figure A. 6 - Particle density regions for the parameters D_{ax} and K_{SEC}

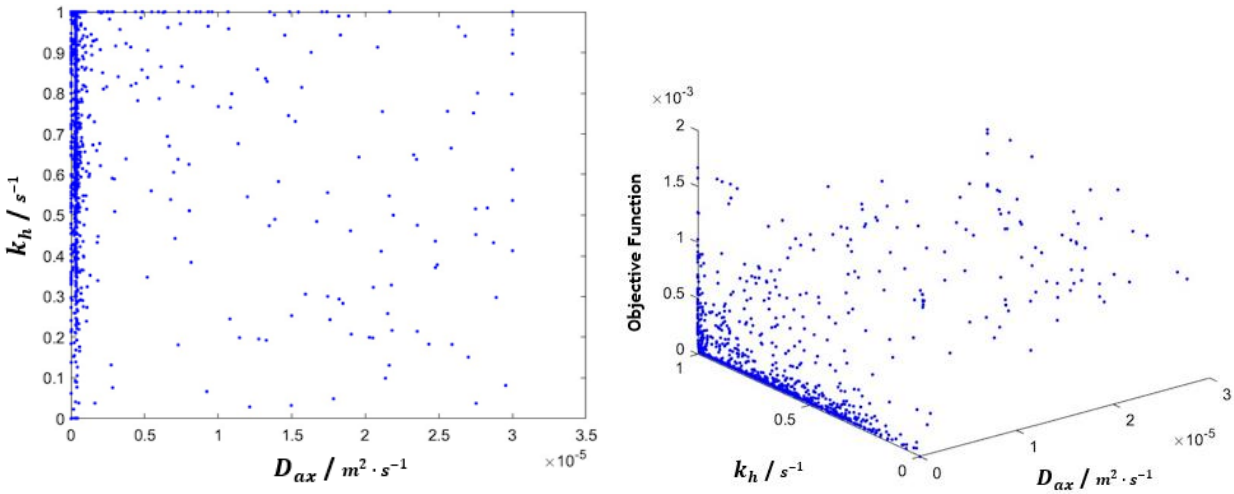

Figure A. 7 - Particle density regions for the parameters D_{ax} and k_h

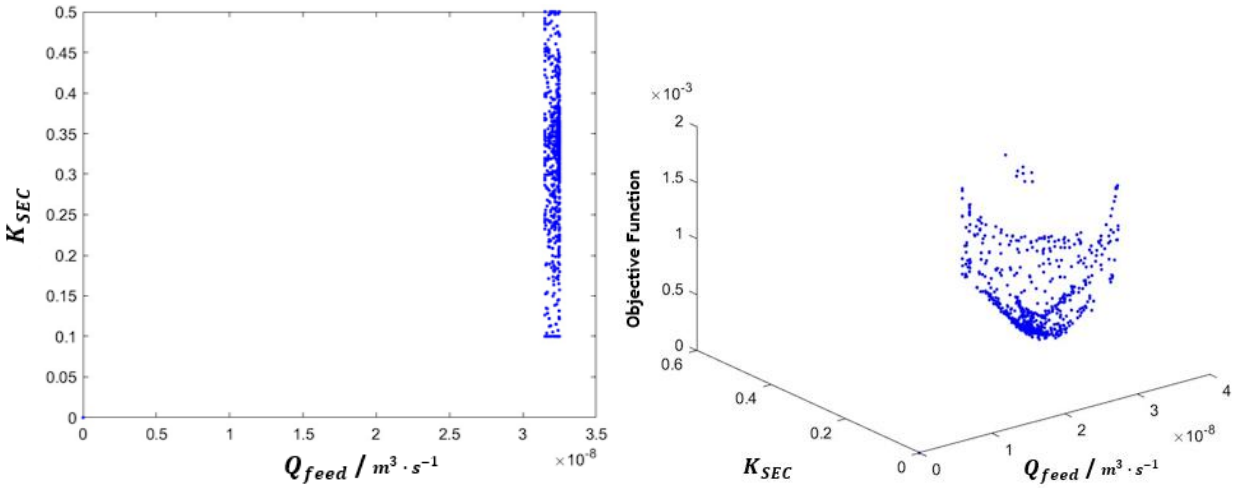

Figure A. 8 - Particle density regions for the parameters Q_{feed} and K_{SEC}

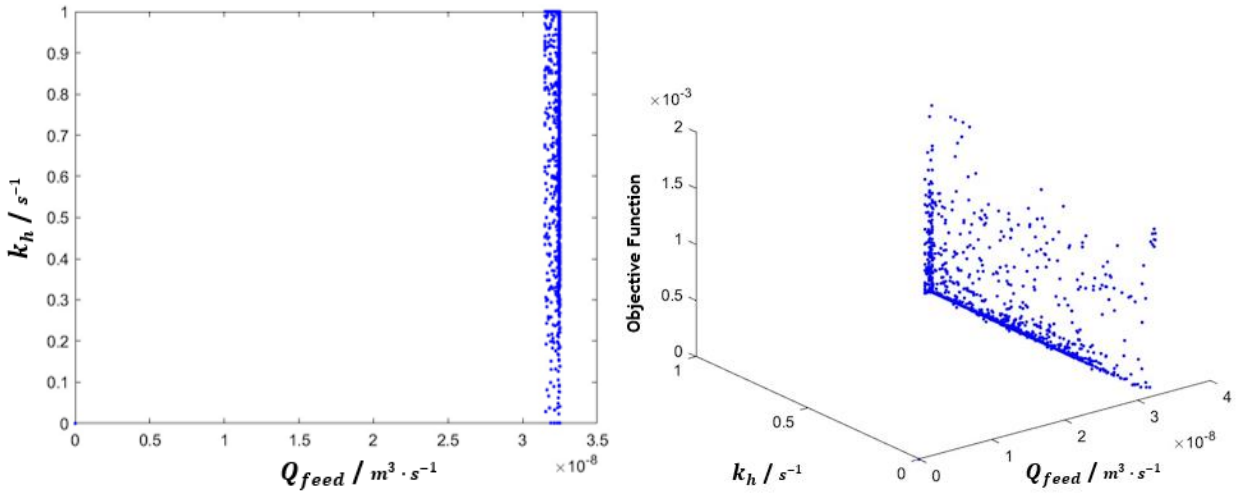

Figure A. 9 - Particle density regions for the parameters Q_{feed} and k_h

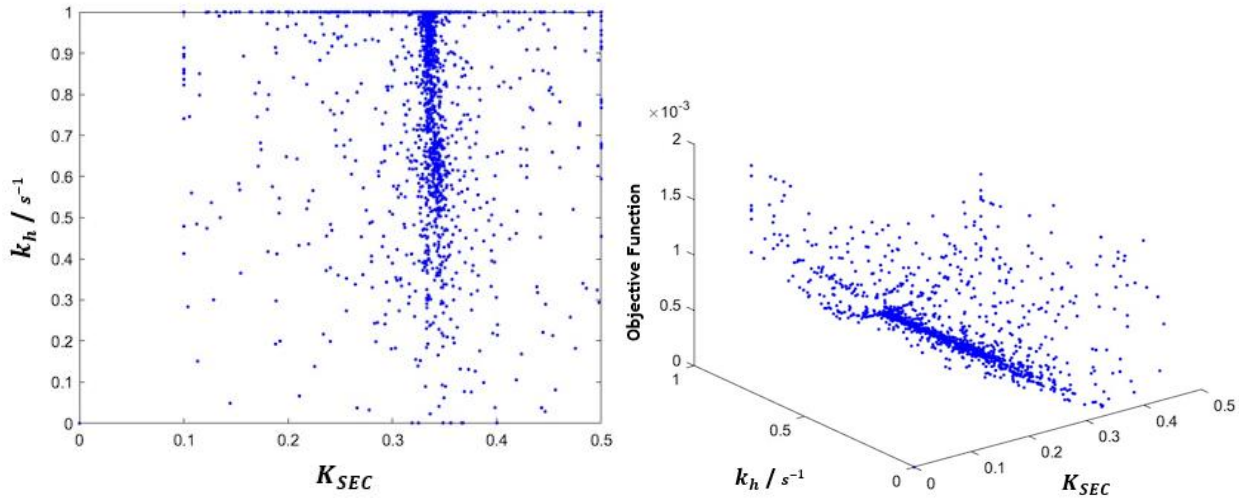

Figure A. 10 - Particle density regions for the parameters K_{SEC} and k_h

A.2. Particle Density Regions for the BSA Experiment:

Similar to the myoglobin experiment, there were 5 parameters chosen to undergo estimation: ε_b, D_{ax}, Q_{feed}, K_{SEC} and k_h. Notably, the K_{SEC} and k_h parameters are different than those in the myoglobin experiment, since these pertain to the BSA protein.

The analysis of the following graphs, aided by the estimability evaluation, led to the removal of one of the parameters in further estimations, k_h. The graphs were also important to facilitate the adjustment of the parameter search region limits for the total parameter estimation. Figure A.11 to Figure A.20 provide information about the positions assumed during the PSO algorithm while also presenting the associated objective function values.

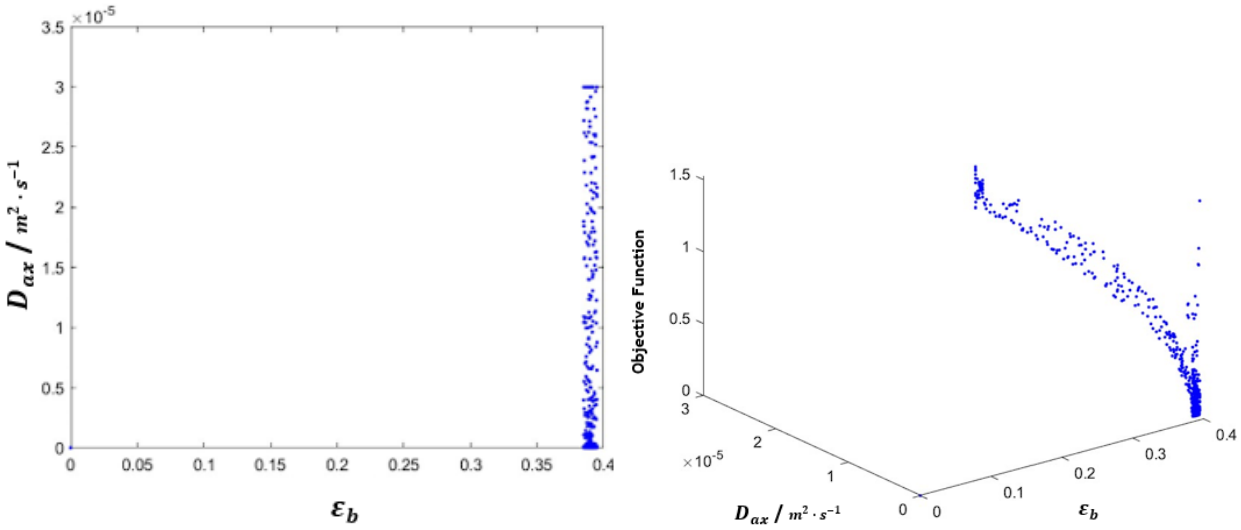

Figure A. 11 - Particle density regions for the parameters ε_b and D_{ax}

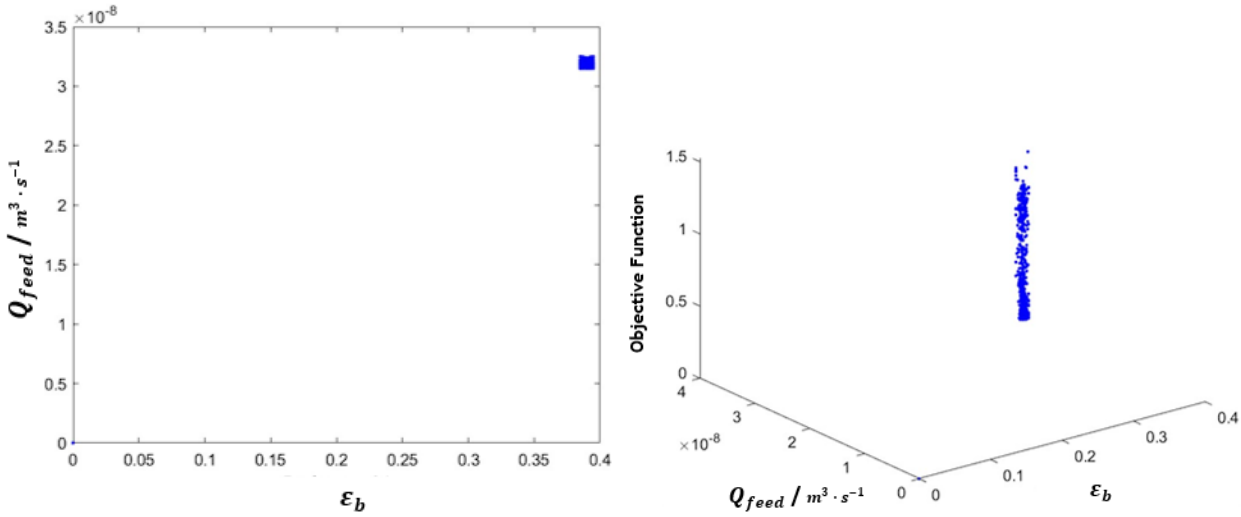

Figure A. 12 - Particle density regions for the parameters ε_b and Q_{feed}

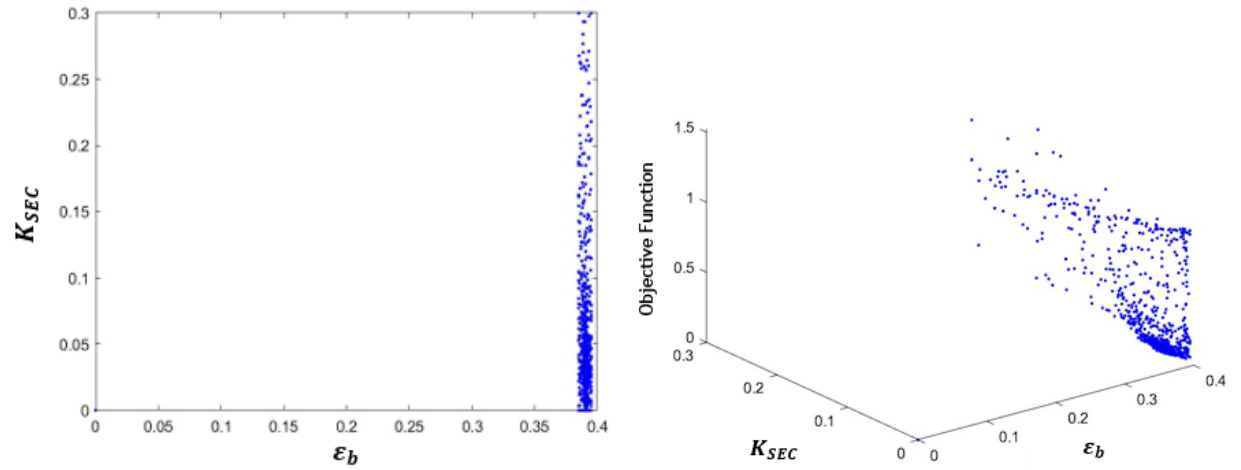

Figure A. 13 - Particle density regions for the parameters ε_b and K_{SEC}

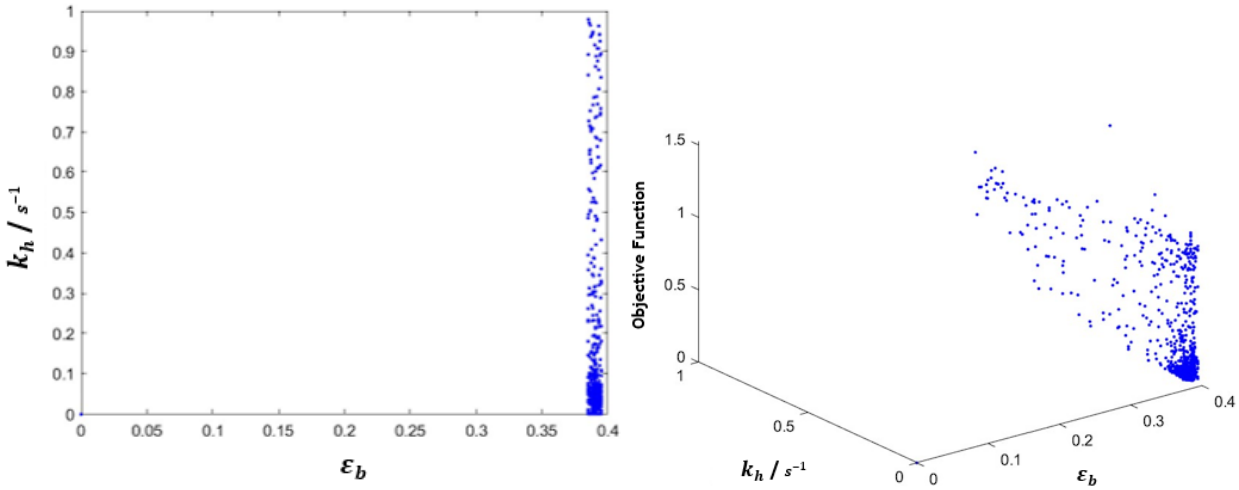

Figure A. 14 - Particle density regions for the parameters ε_b and k_h

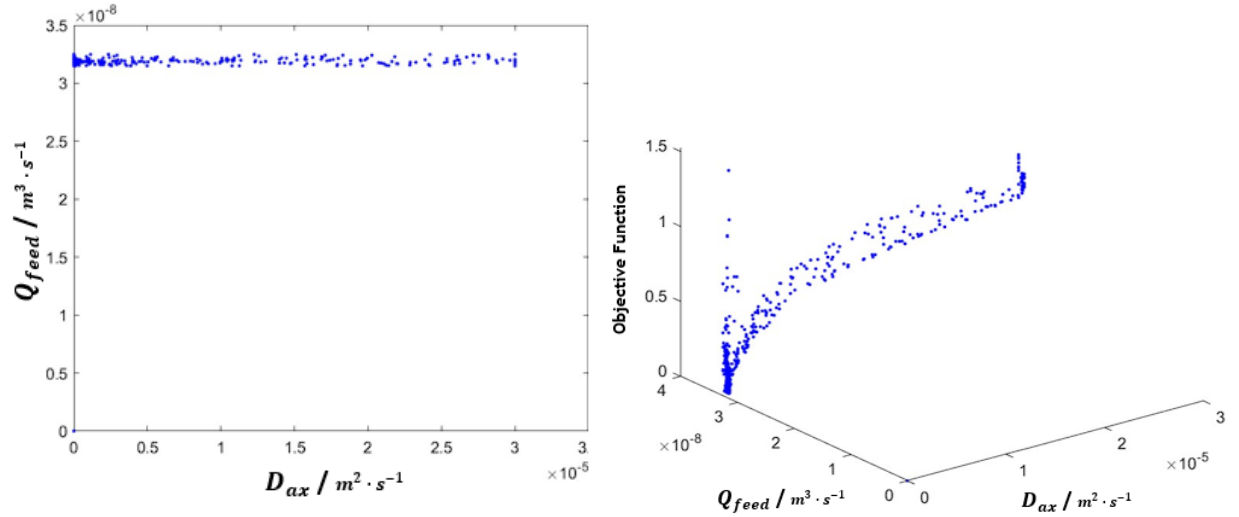

Figure A. 15 - Particle density regions for the parameters D_{ax} and Q_{feed}

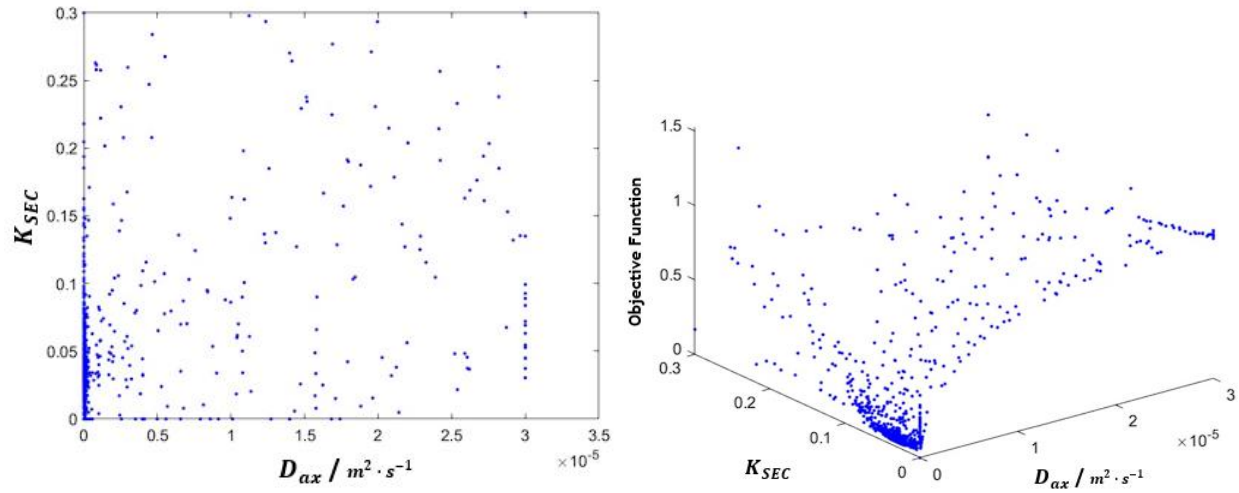

Figure A. 16 - Particle density regions for the parameters D_{ax} and K_{SEC}

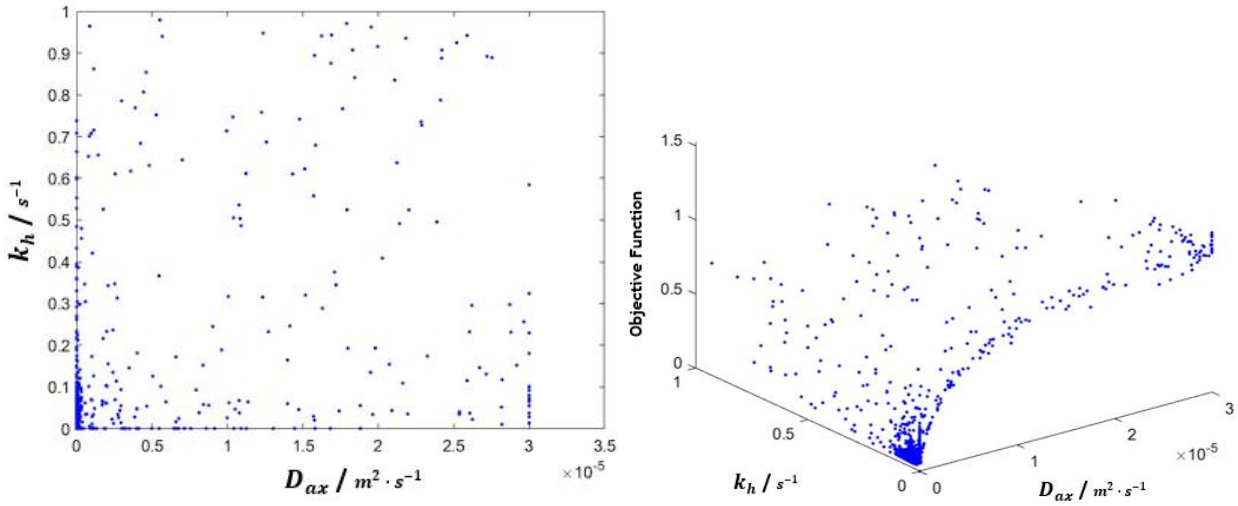

Figure A. 17 - Particle density regions for the parameters D_{ax} and k_h

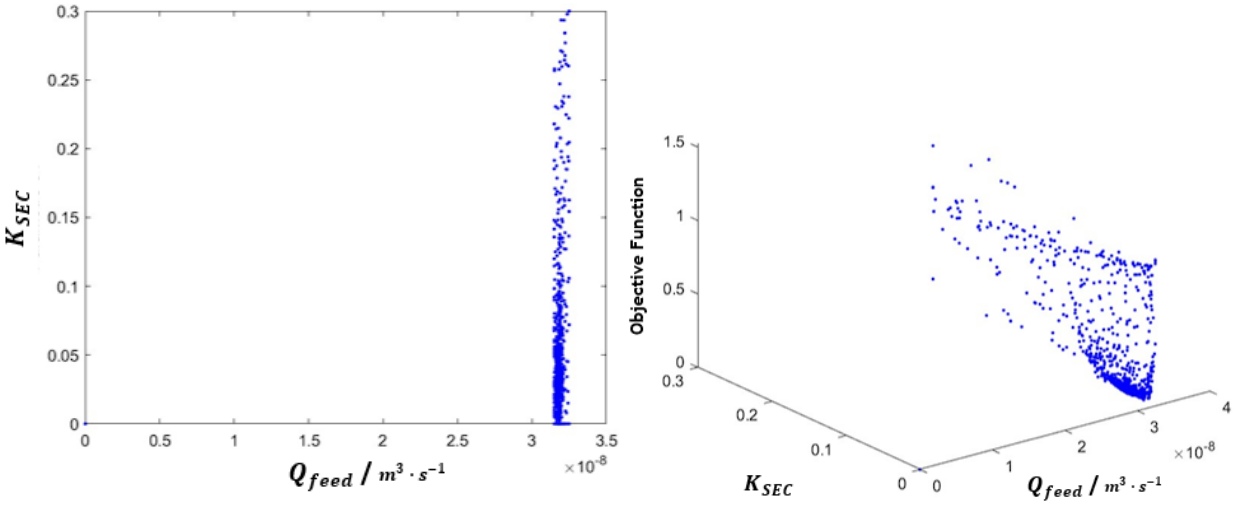

Figure A. 18 - Particle density regions for the parameters Q_{feed} and K_{SEC}

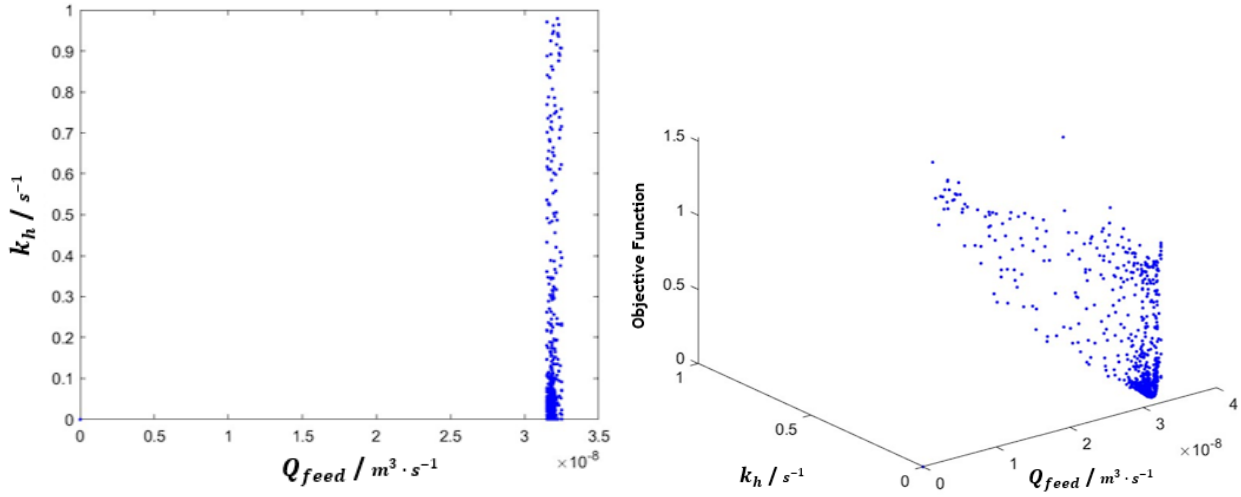

Figure A. 19 - Particle density regions for the parameters Q_{feed} and k_h

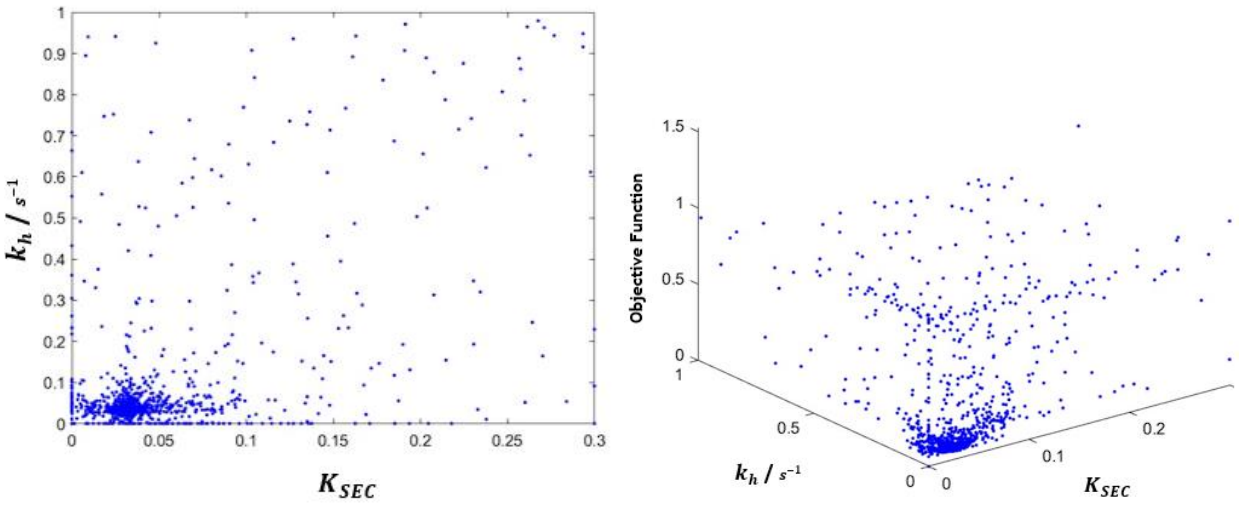

Figure A. 20 - Particle density regions for the parameters K_{SEC} and k_h

A.3. Particle Density Regions for the Binary Experiment:

For the binary fixed-bed experiment, there were 7 parameters selected to be estimated: ε_b, D_{ax}, Q_{feed}, $K_{SEC,BSA}$, $K_{SEC,Mb}$, $k_{h,BSA}$ and $k_{h,Mb}$.

The observation of the graphs shown next, aided by the estimability evaluation, led to the removal of two of the parameters in further estimations, the mass transfer coefficients, $k_{h,BSA}$ and $k_{h,Mb}$. Once more, the graphs were important to ease the adjustment of the parameter search region limits for the total parameter estimation. Figure A.21 to Figure A.41 provide information about the particles' positions yielded by the PSO algorithm while also presenting the objective function values associated with each particle.

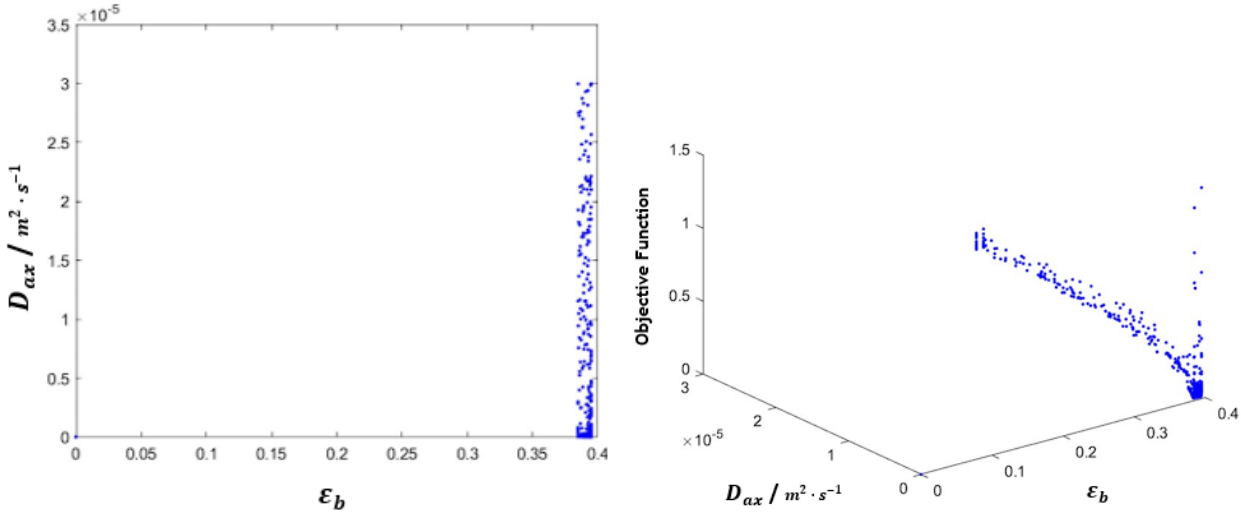

Figure A. 21 - Particle density regions for the parameters ε_b and D_{ax}

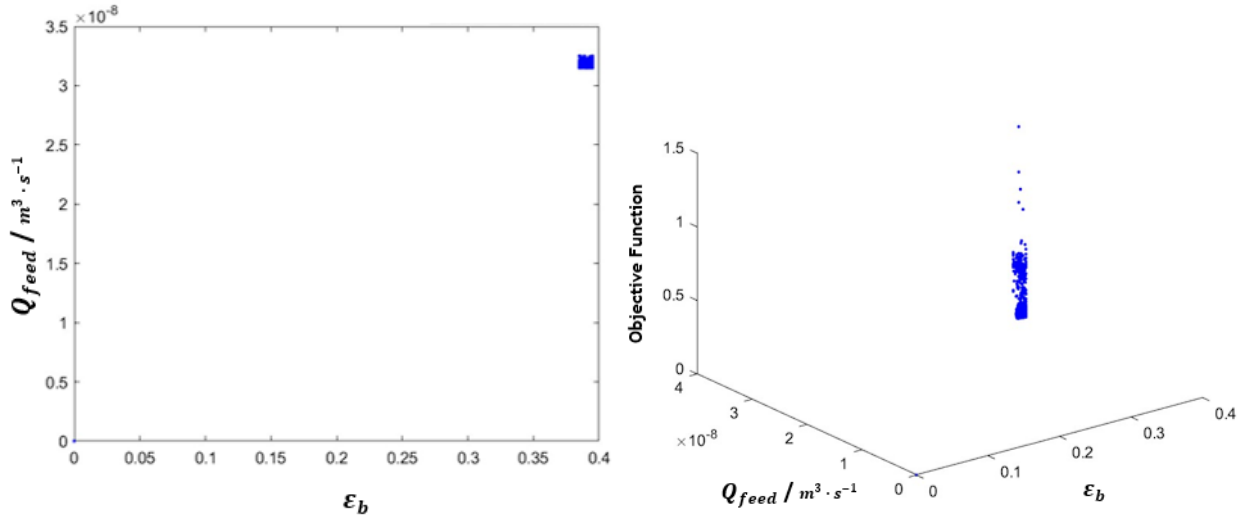

Figure A. 22 - Particle density regions for the parameters ε_b and Q_{feed}

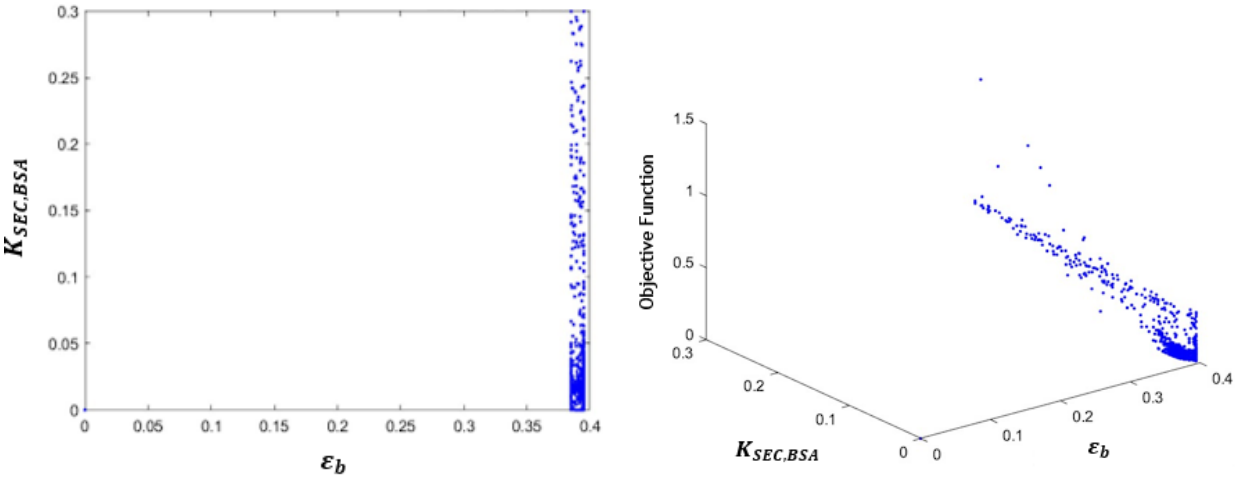

Figure A. 23 - Particle density regions for the parameters ε_b and $K_{SEC,BSA}$

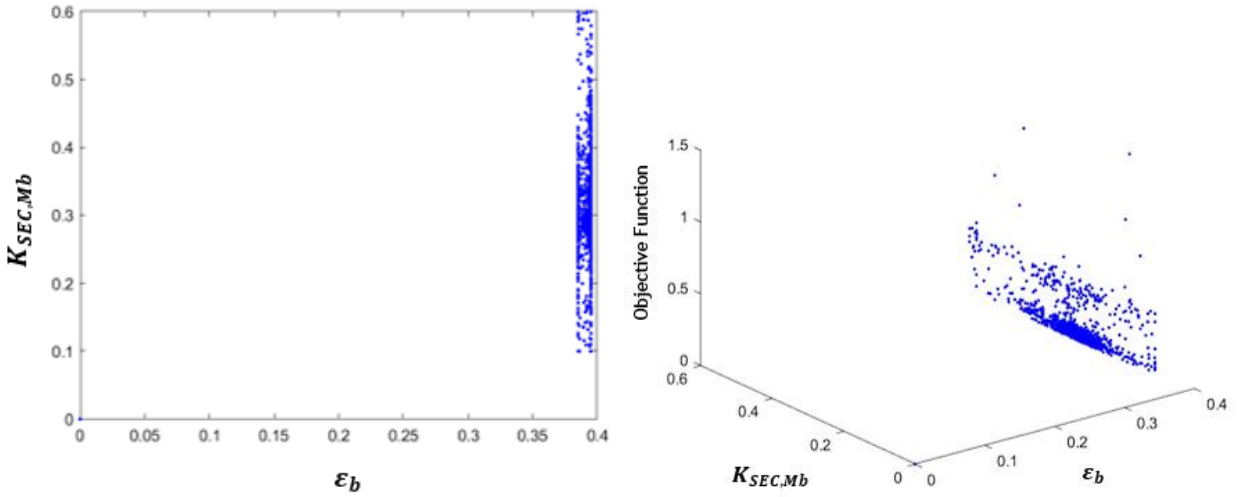

Figure A. 24 - Particle density regions for the parameters ε_b and $K_{SEC,Mb}$

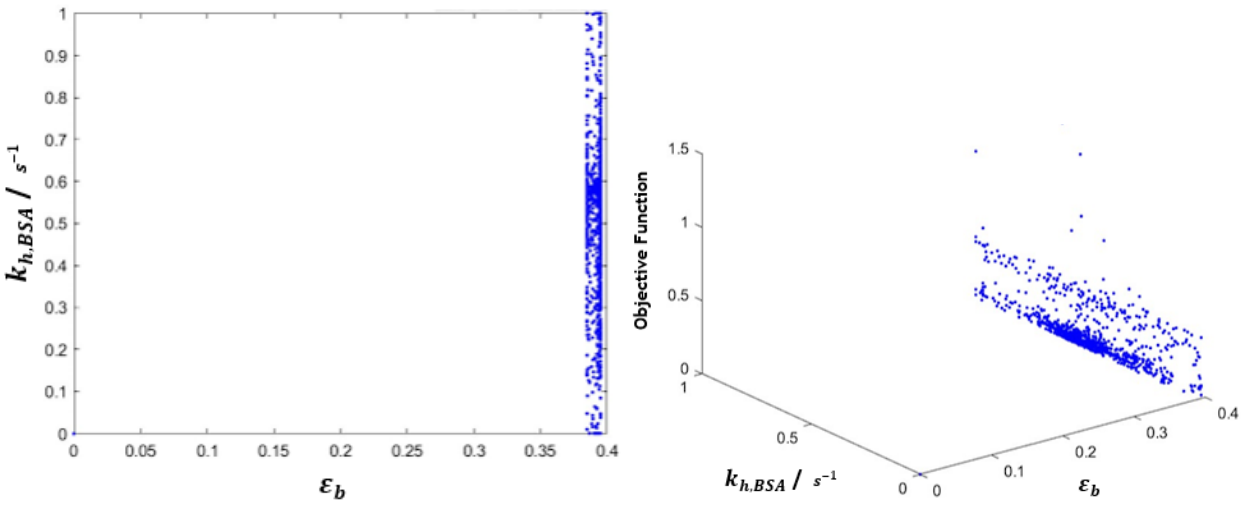

Figure A. 25 - Particle density regions for the parameters ε_b and $k_{h,BSA}$

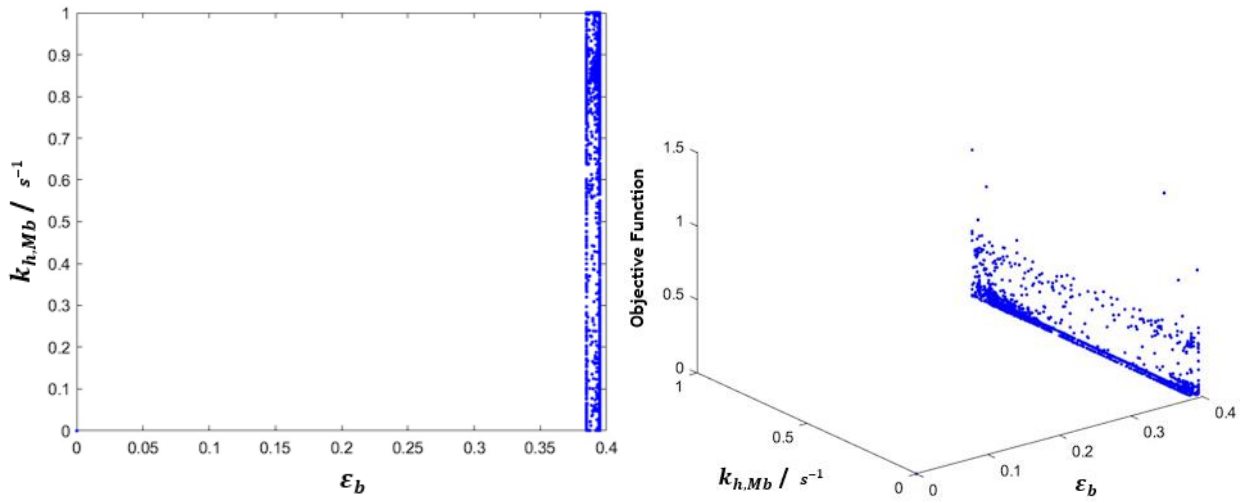

Figure A. 26 - Particle density regions for the parameters ε_b and $k_{h,Mb}$

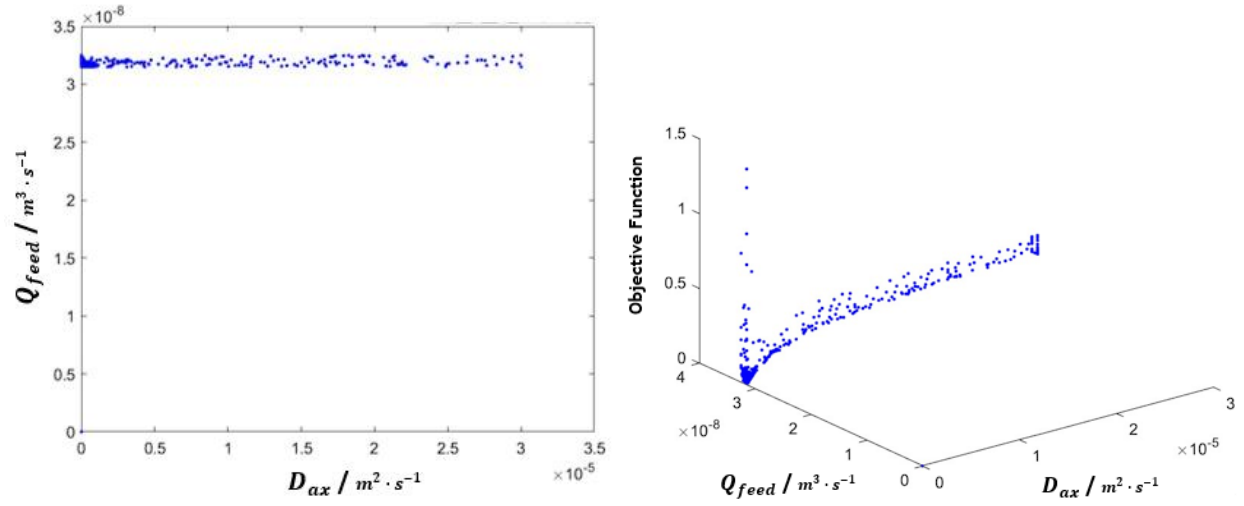

Figure A. 27 - Particle density regions for the parameters D_{ax} and Q_{feed}

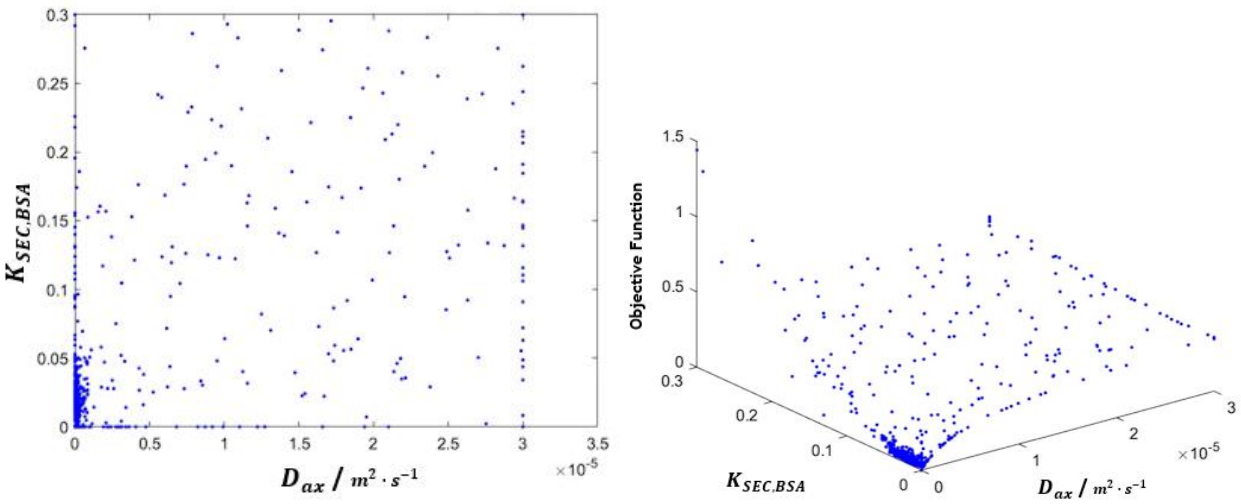

Figure A. 28 - Particle density regions for the parameters D_{ax} and $K_{SEC,BSA}$

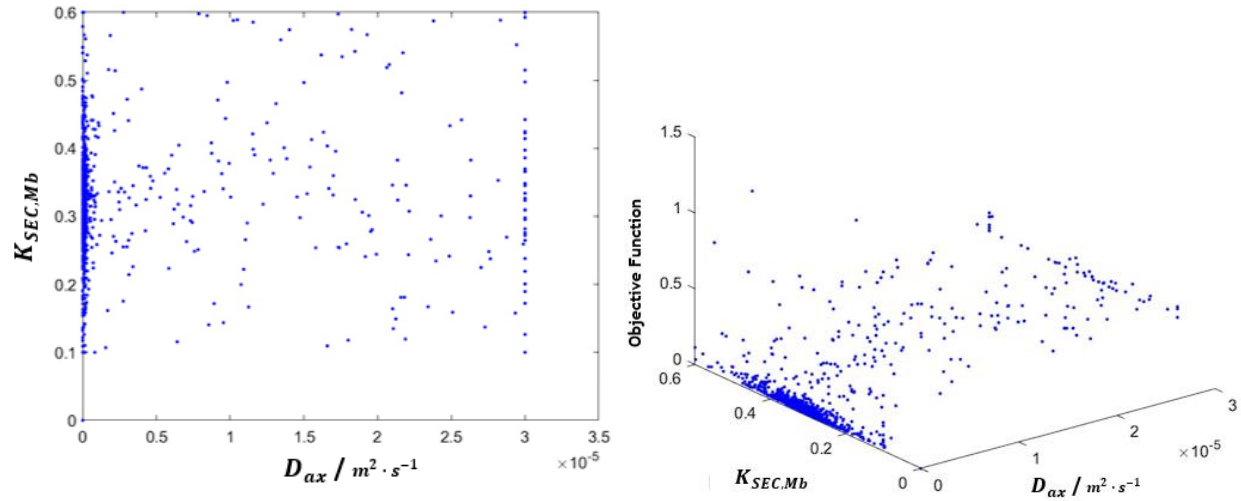

Figure A. 29 - Particle density regions for the parameters D_{ax} and $K_{SEC,Mb}$

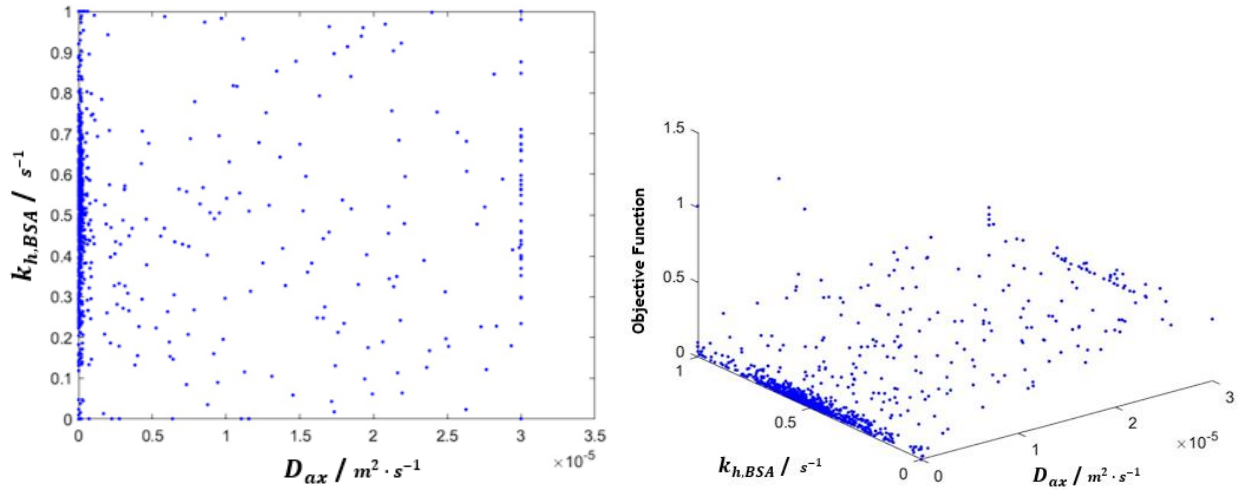

Figure A. 30 - Particle density regions for the parameters D_{ax} and $k_{h,BSA}$

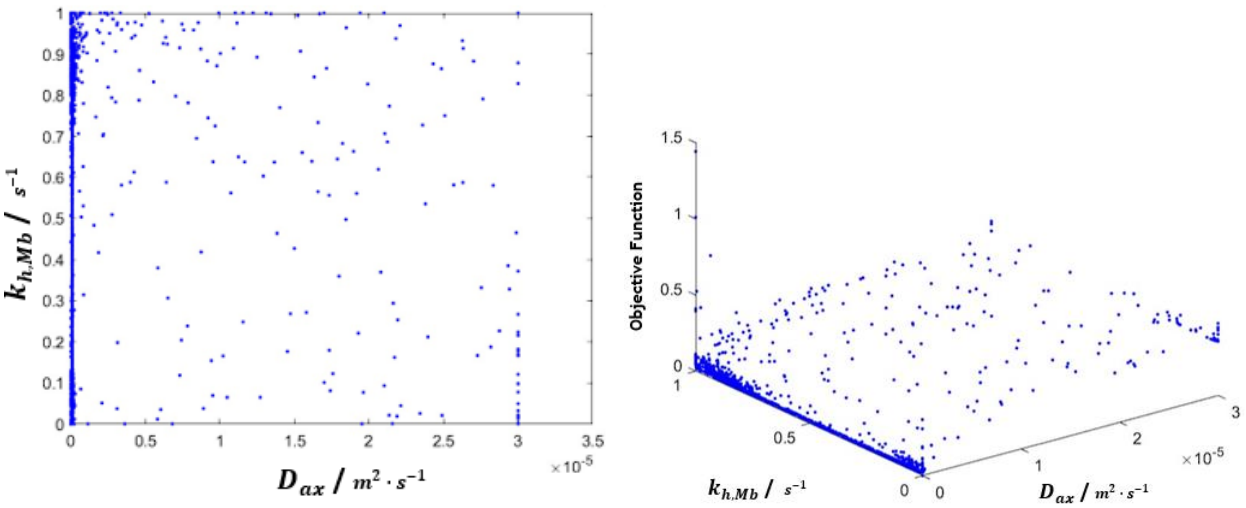

Figure A. 31 - Particle density regions for the parameters D_{ax} and $k_{h,Mb}$

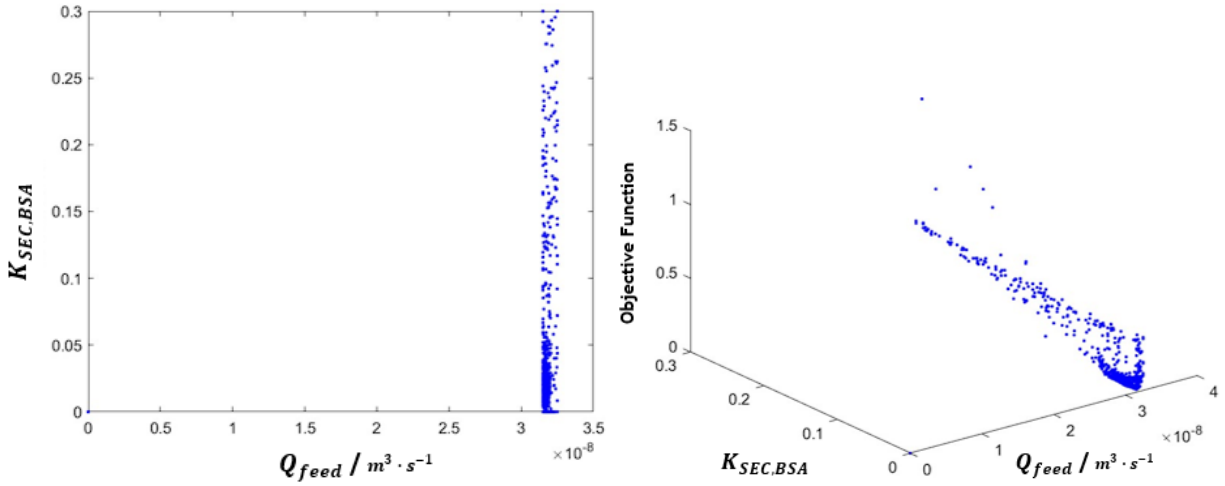

Figure A. 32 - Particle density regions for the parameters Q_{feed} and $K_{SEC,BSA}$

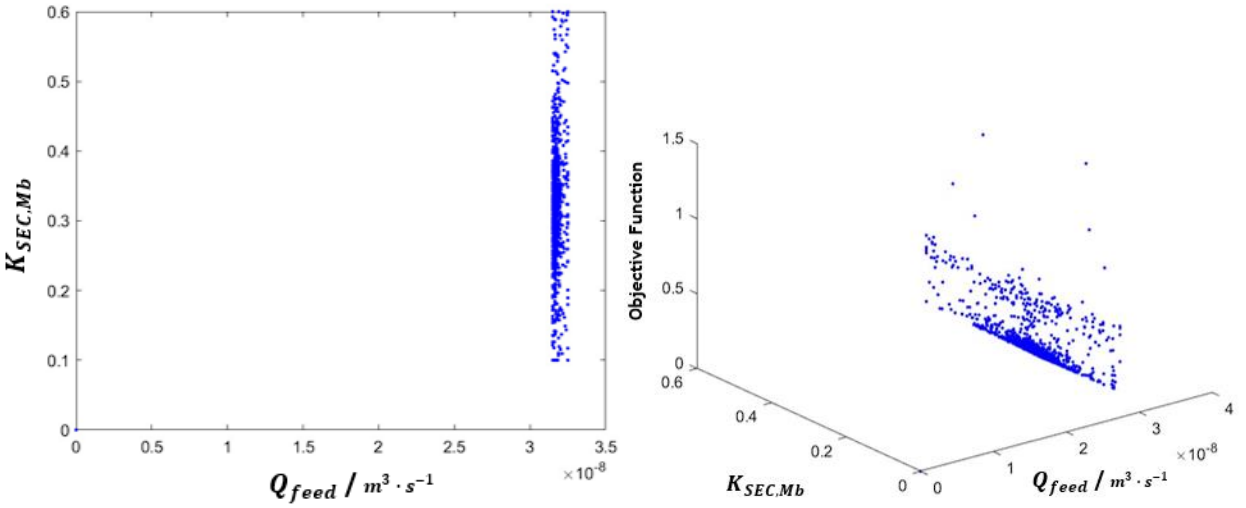

Figure A. 33 - Particle density regions for the parameters Q_{feed} and $K_{SEC,Mb}$

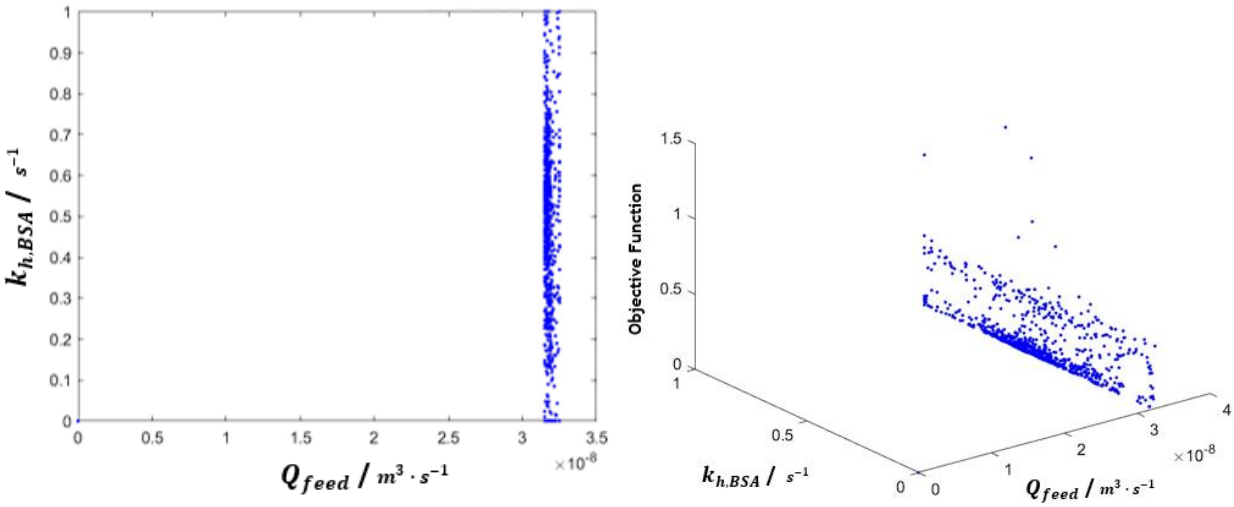

Figure A. 34 - Particle density regions for the parameters Q_{feed} and $k_{h,BSA}$

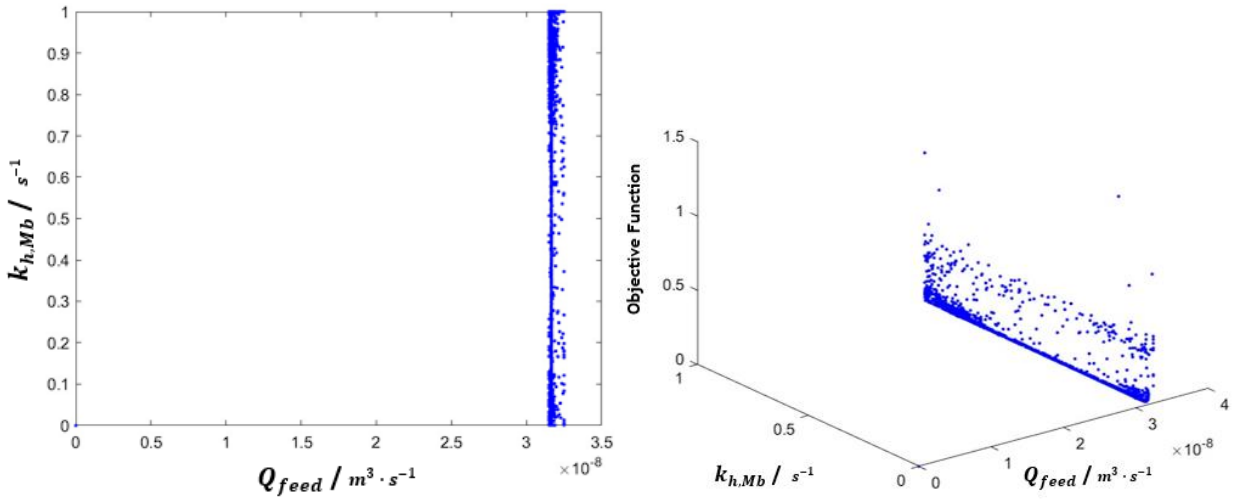

Figure A. 35 - Particle density regions for the parameters Q_{feed} and $k_{h,Mb}$

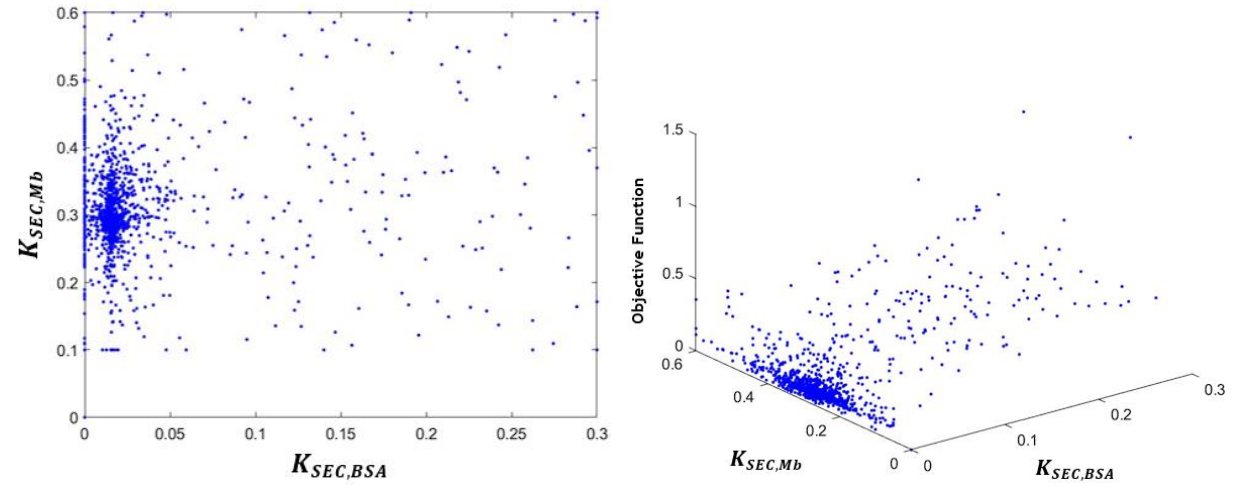

Figure A. 36 - Particle density regions for the parameters $K_{SEC,BSA}$ and $K_{SEC,Mb}$

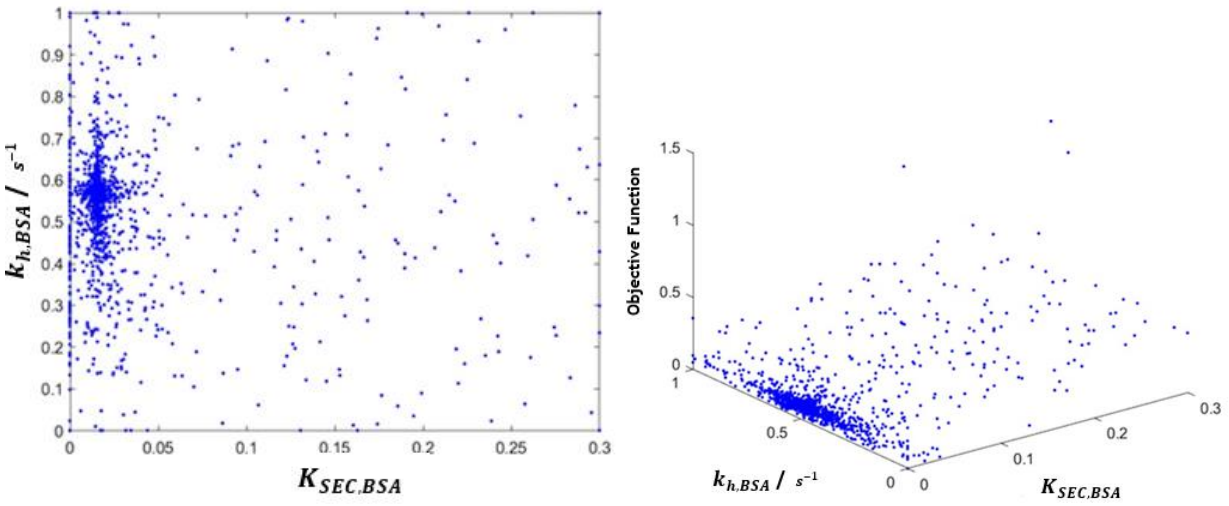

Figure A. 37 - Particle density regions for the parameters $K_{SEC,BSA}$ and $k_{h,BSA}$

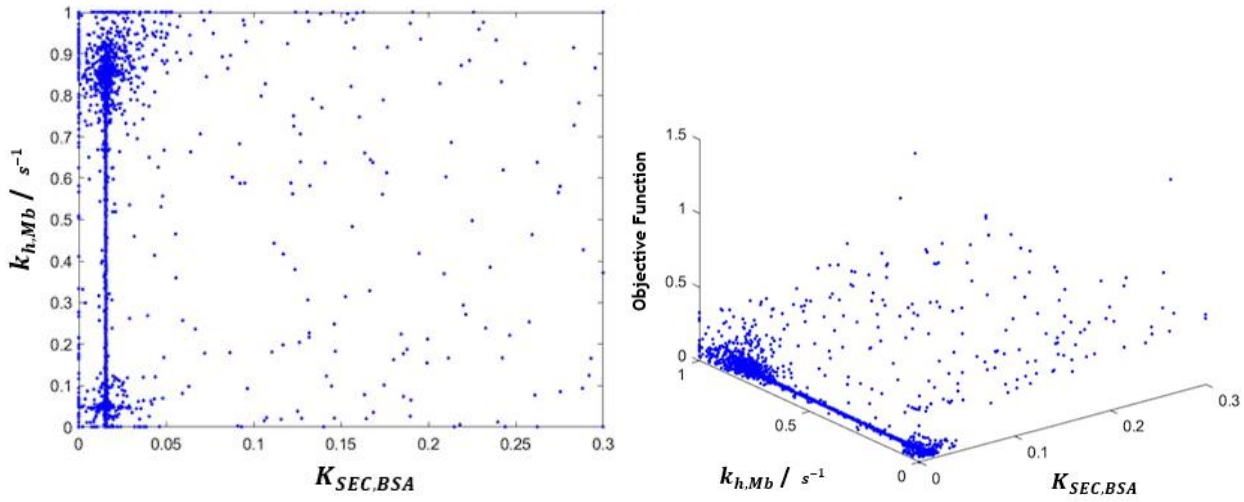

Figure A. 38 - Particle density regions for the parameters $K_{SEC,BSA}$ and $k_{h,Mb}$

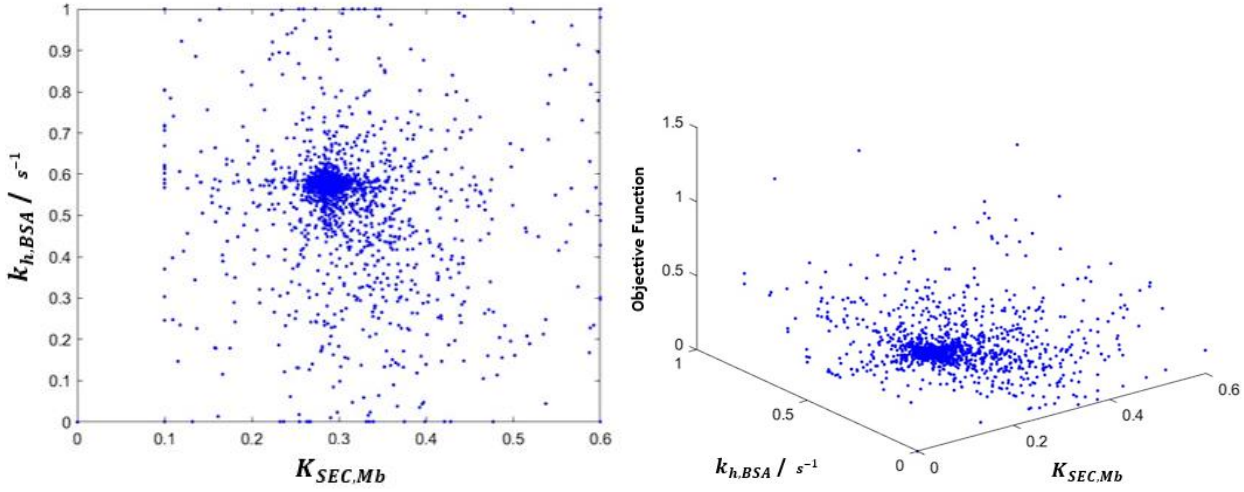

Figure A. 39 - Particle density regions for the parameters $K_{SEC,Mb}$ and $k_{h,BSA}$

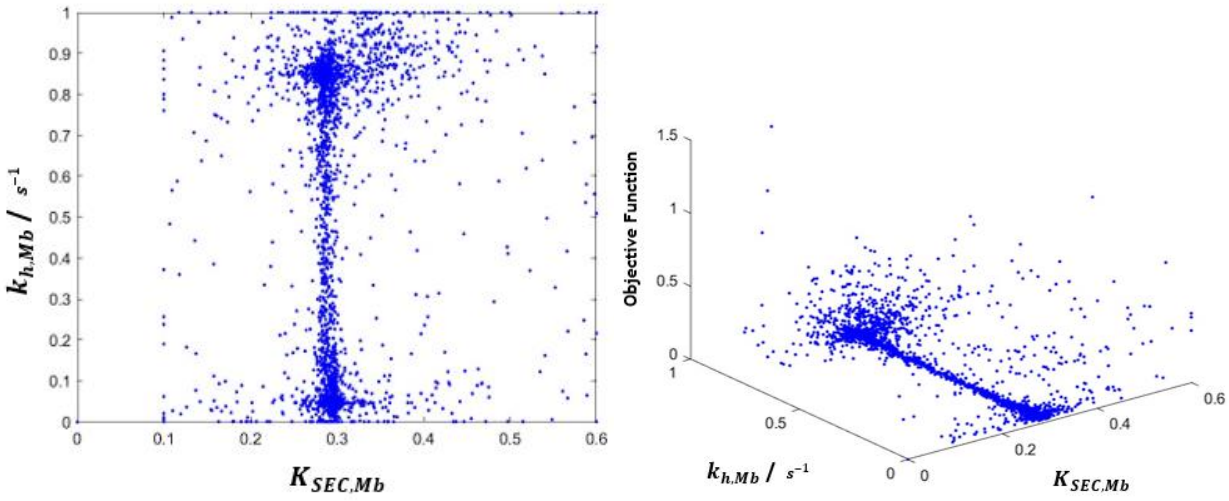

Figure A. 40 - Particle density regions for the parameters $K_{SEC,Mb}$ and $k_{h,Mb}$

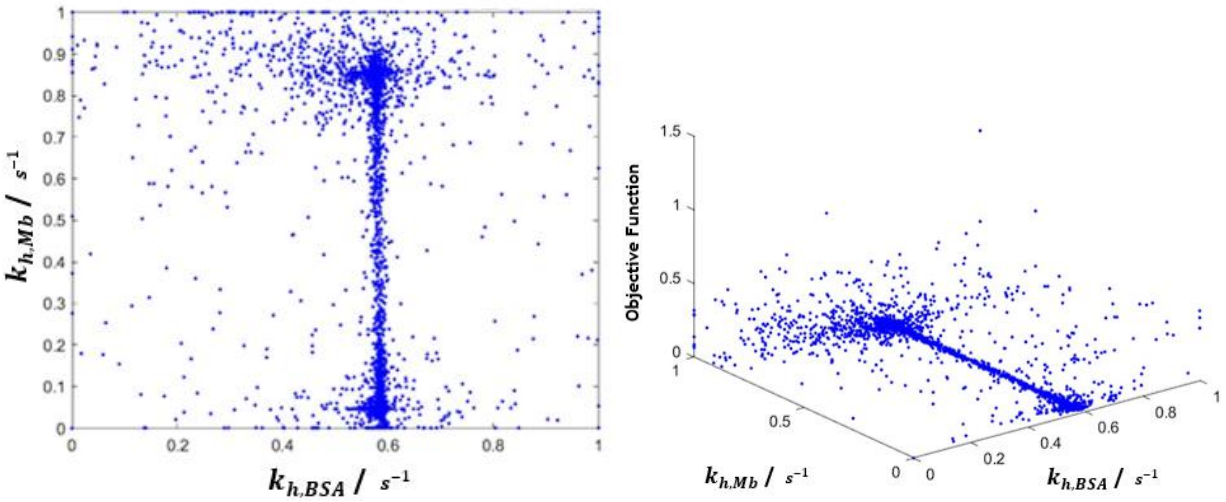

Figure A. 41 - Particle density regions for the parameters $k_{h,BSA}$ and $k_{h,Mb}$

Appendix B - Confidence Regions

In the following pages, the confidence regions of twenty-one estimable parameters, with a confidence of 95%, will be presented. These regions were built based on the Fisher-Snedecor test and used the determined uncertainty of each parameter.

The axes of the graphs are limited by the defined boundaries of each parameter, where the lower boundary marks the beginning of the axis and the upper boundary represents the farthest point. The confidence regions are presented by the clouds of blue particles and the ellipse represents the confidence region if the normal distribution condition was verified; therefore, it works as a comparison to the likelihood region yielded by the Fisher-Snedecor test. The mean value obtained from the particles during the total parameter estimation is shown by a black cross, and the best value obtained from the Particle Swarm Optimization (PSO) is revealed by a red cross. As a reminder, the axial dispersion coefficient, $D_{ax,}$ is given in $m^2 \cdot s^{-1}$, the fixed bed feed flowrate, Q_{feed}, is given in $m^3 \cdot s^{-1}$, and lastly, the SMB operational flowrates, Q_F, Q_D,Q_E, and Q_{IV}, are given in $l \cdot \ min^{-1}$. The parameters not mentioned are dimensionless.

Although these graphical representations were not the primary focus of the dissertation, they proved to be valuable tools for understanding and fully characterizing the achieved results.

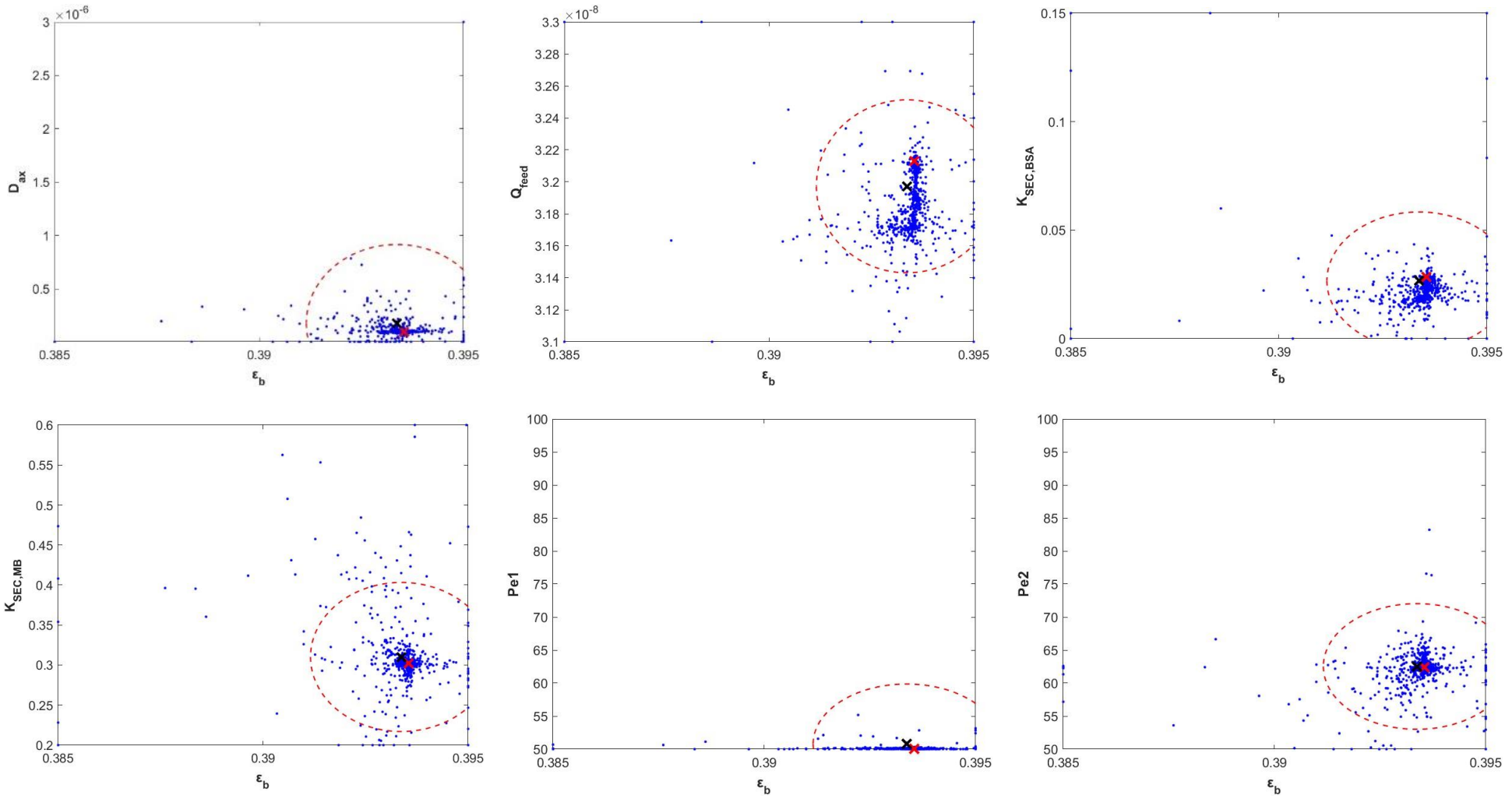

Figure B. 1 – Confidence Regions for each pair of estimable parameters – Pairs containing ε_b

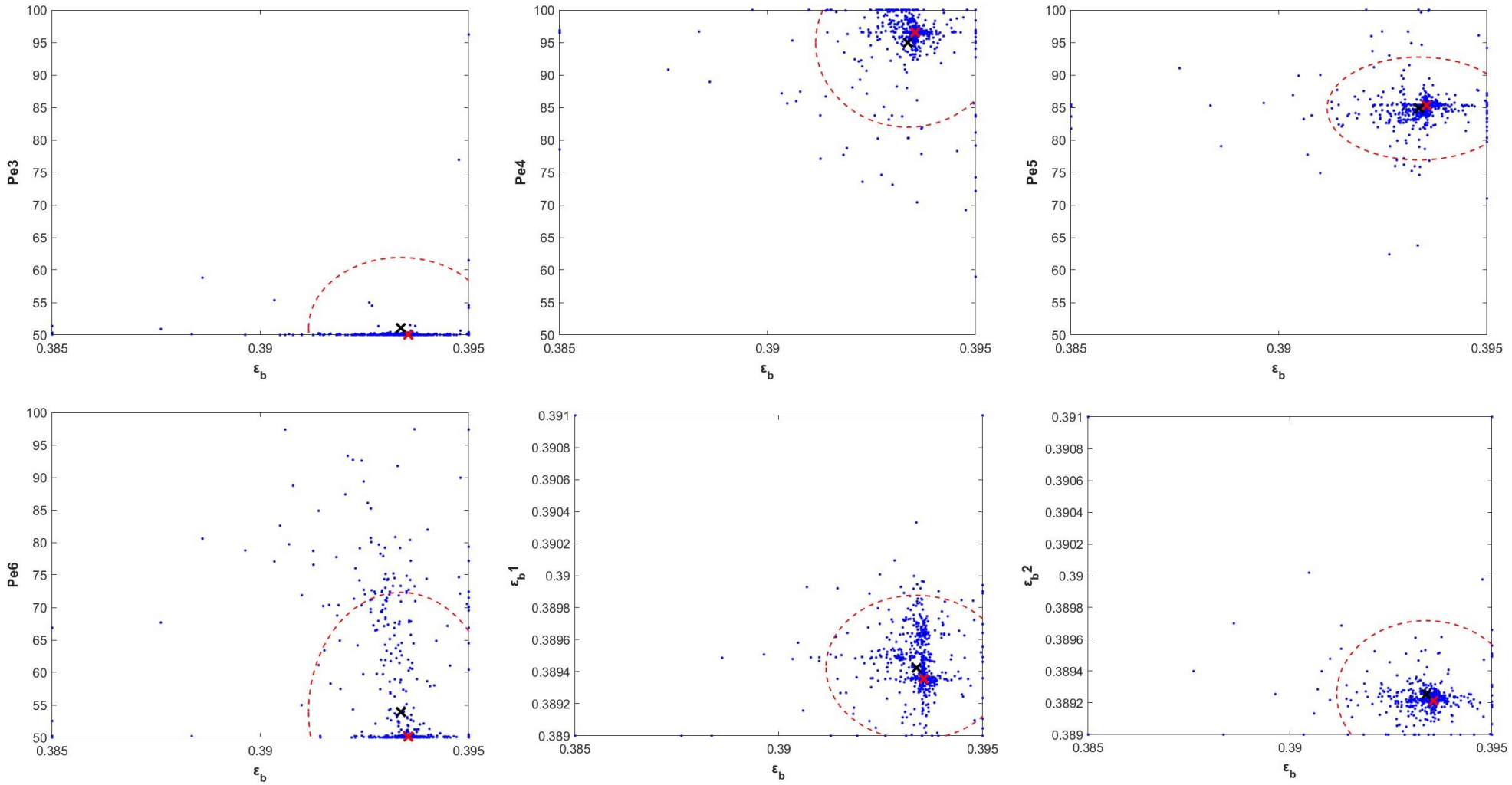

Figure B. 2 - Confidence Regions for each pair of estimable parameters – Pairs containing ε_b (Continuation)

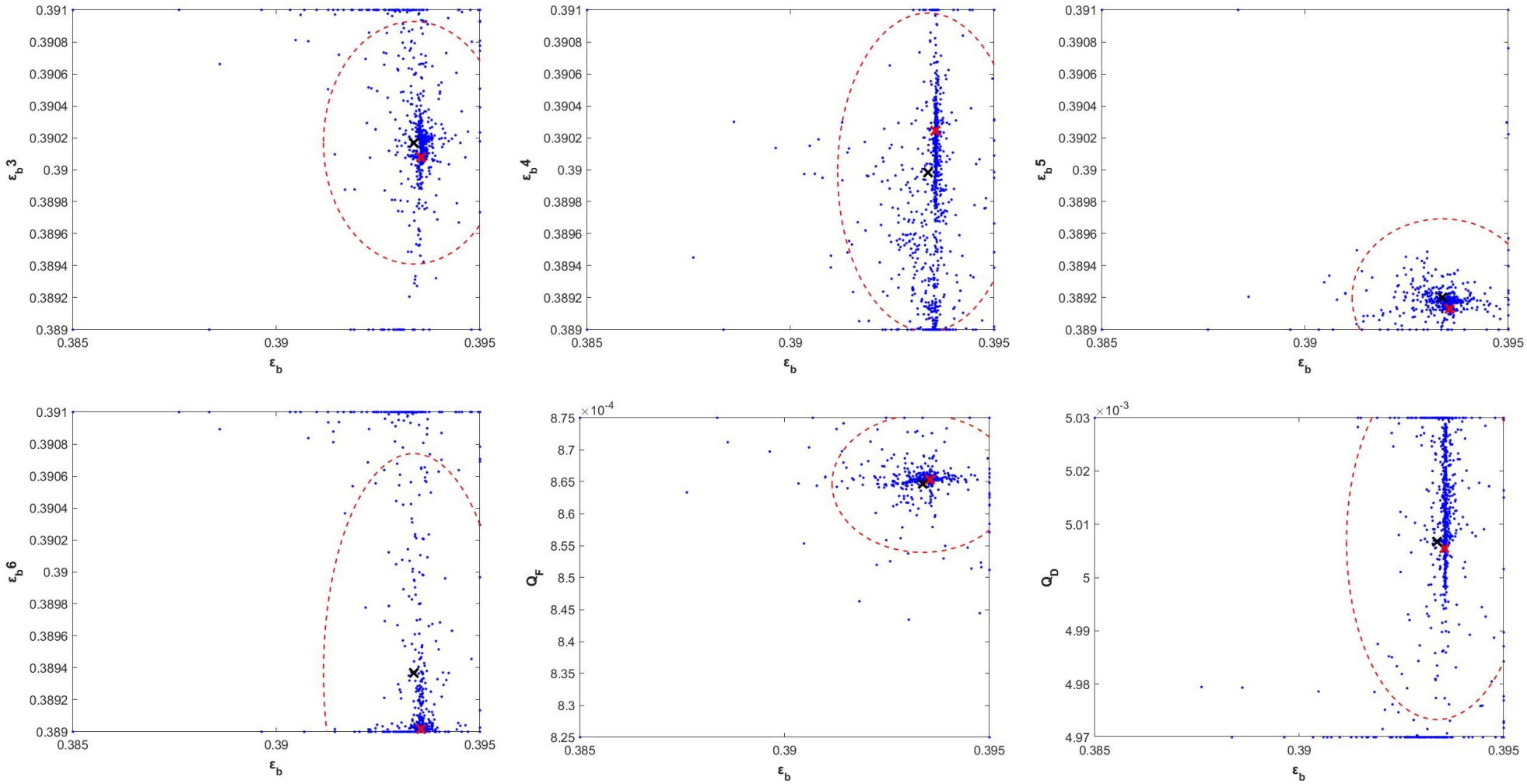

Figure B. 3 - Confidence Regions for each pair of estimable parameters – Pairs containing ε_b (Continuation)

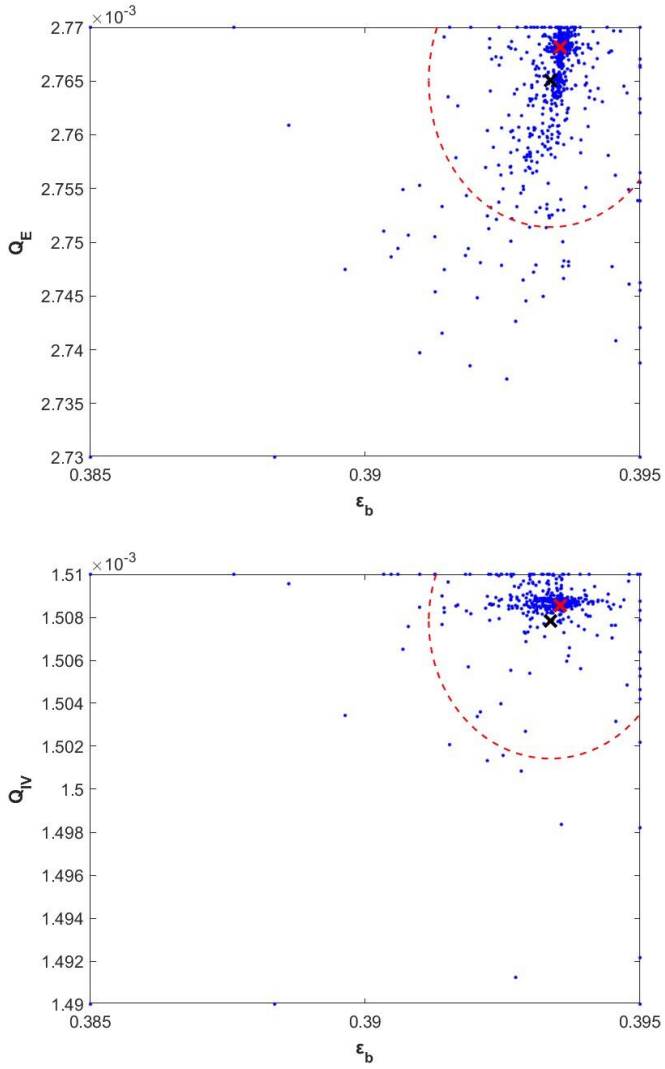

Figure B. 4 - Confidence Regions for each pair of estimable parameters – Pairs containing ε_b (Continuation)

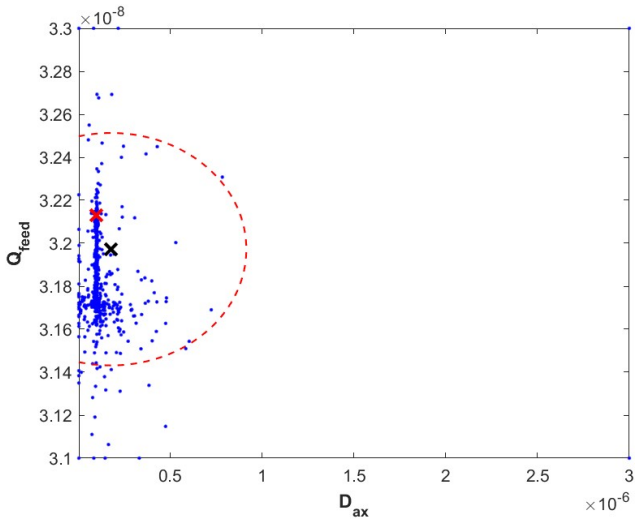

Figure B. 5 - Confidence Regions for each pair of estimable parameters – Pairs containing D_{ax}

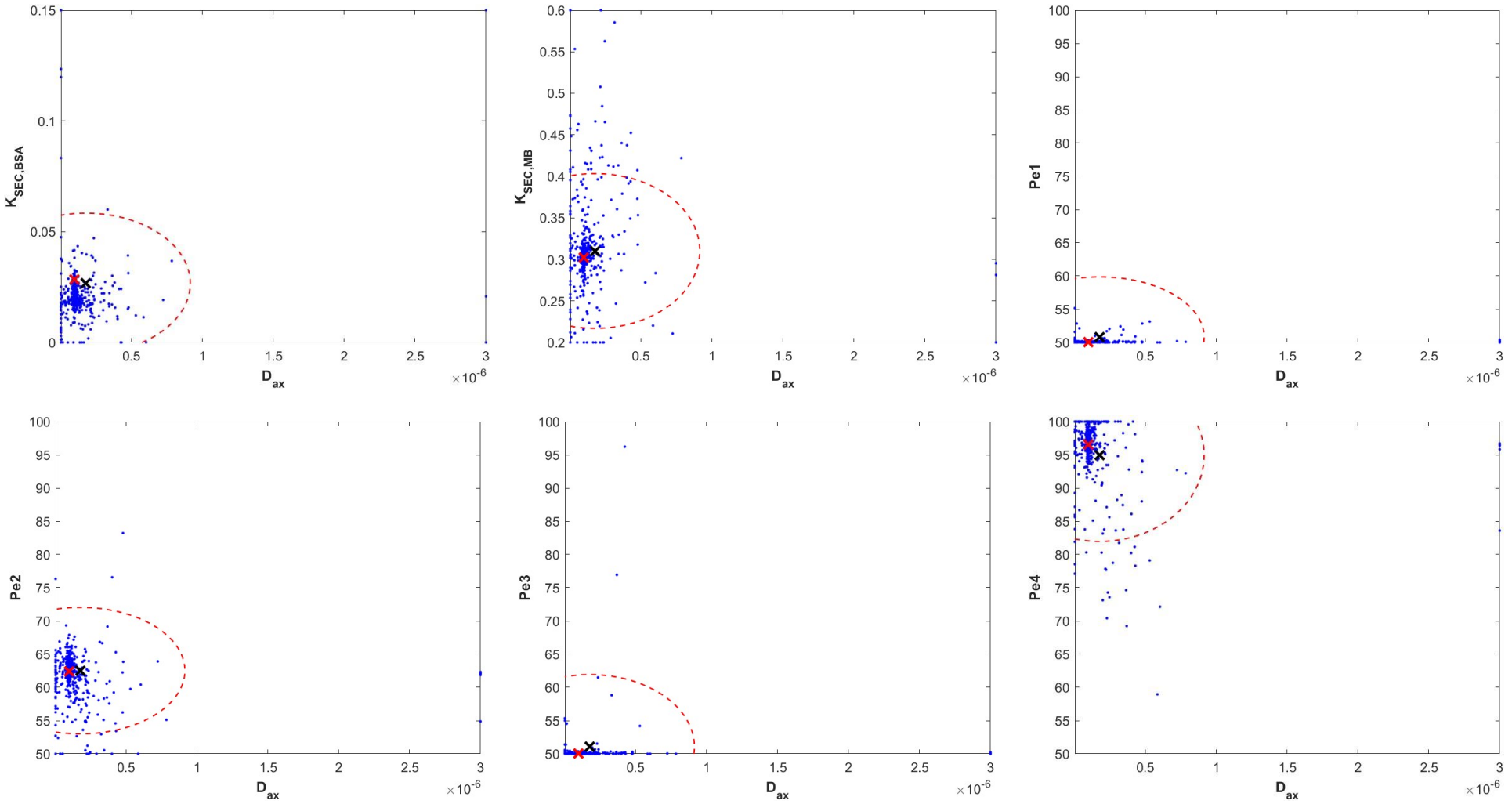

Figure B. 6 - Confidence Regions for each pair of estimable parameters – Pairs containing D_{ax} (Continuation)

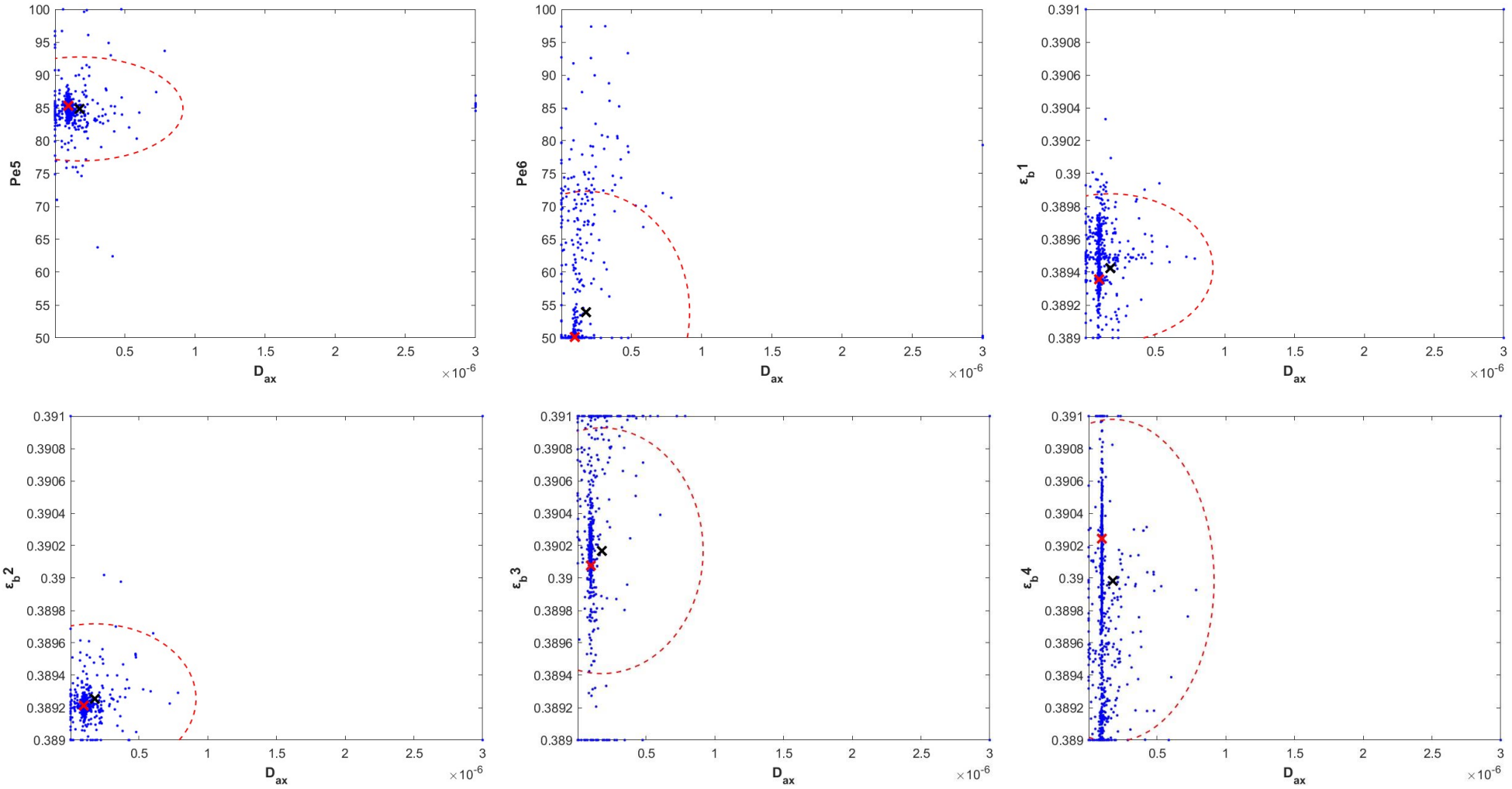

Figure B. 7 - Confidence Regions for each pair of estimable parameters – Pairs containing D_{ax} (Continuation)

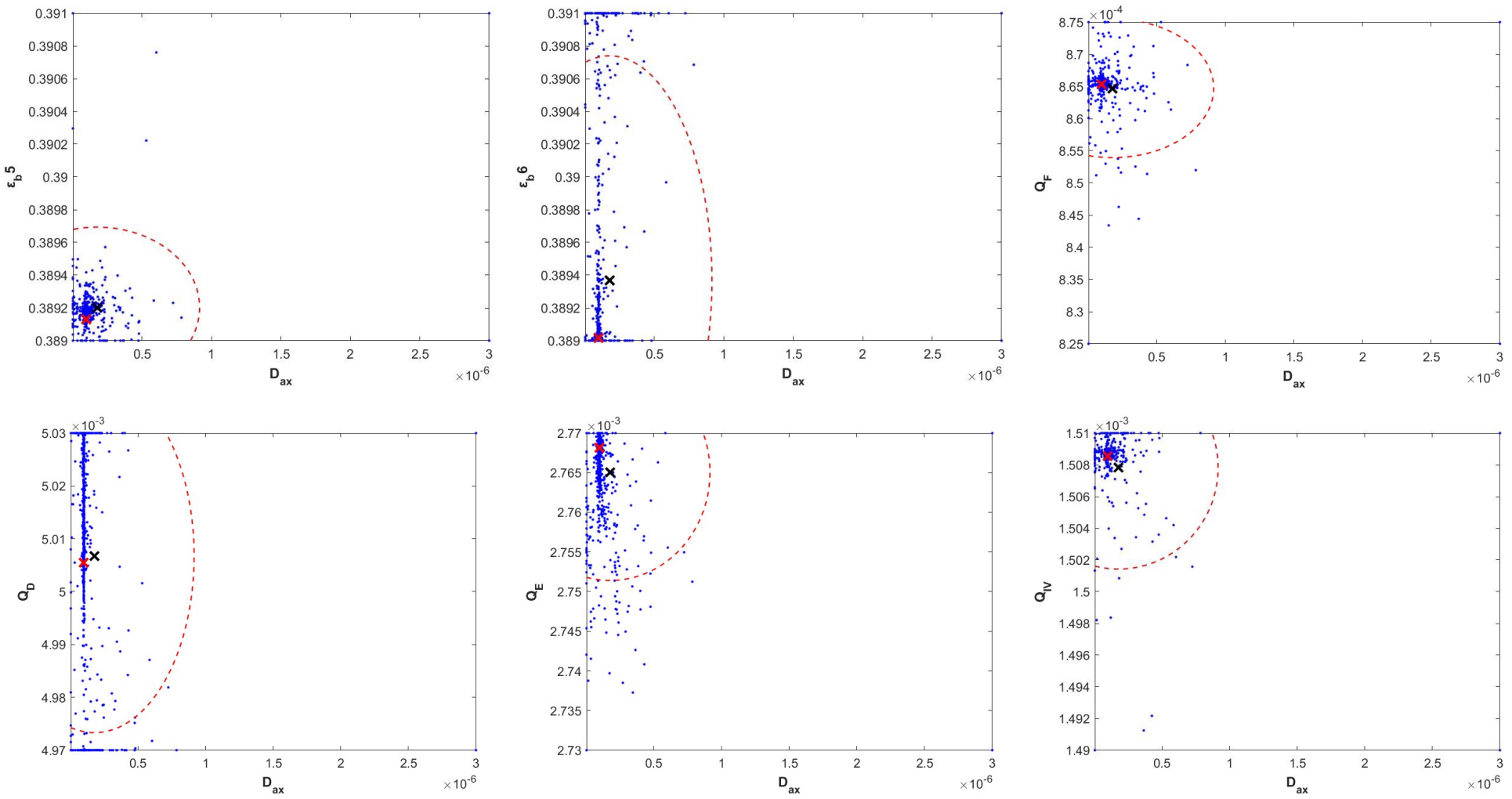

Figure B. 8 - Confidence Regions for each pair of estimable parameters – Pairs containing D_{ax} (Continuation)

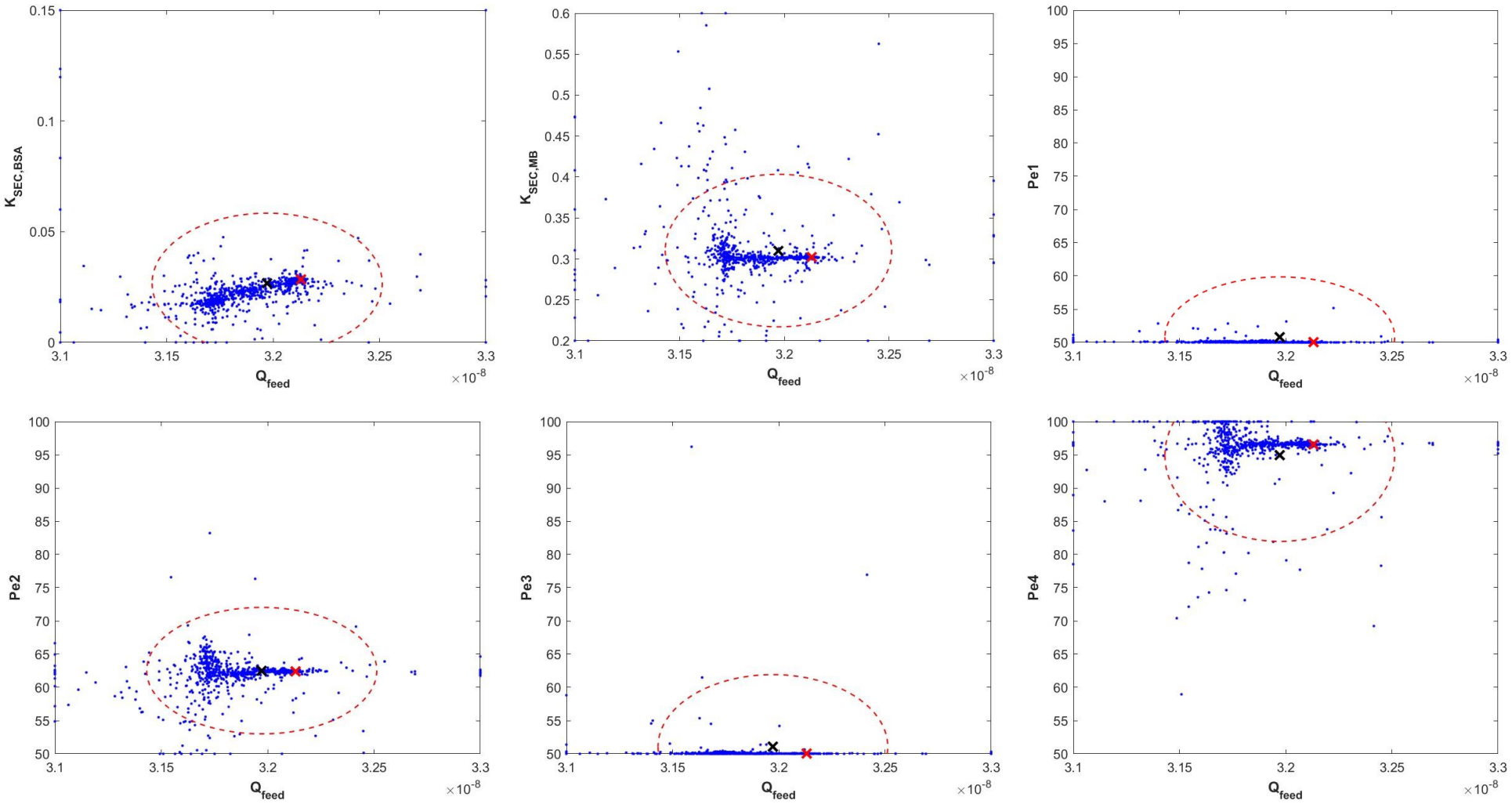

Figure B. 9 - Confidence Regions for each pair of estimable parameters – Pairs containing Q_{feed}

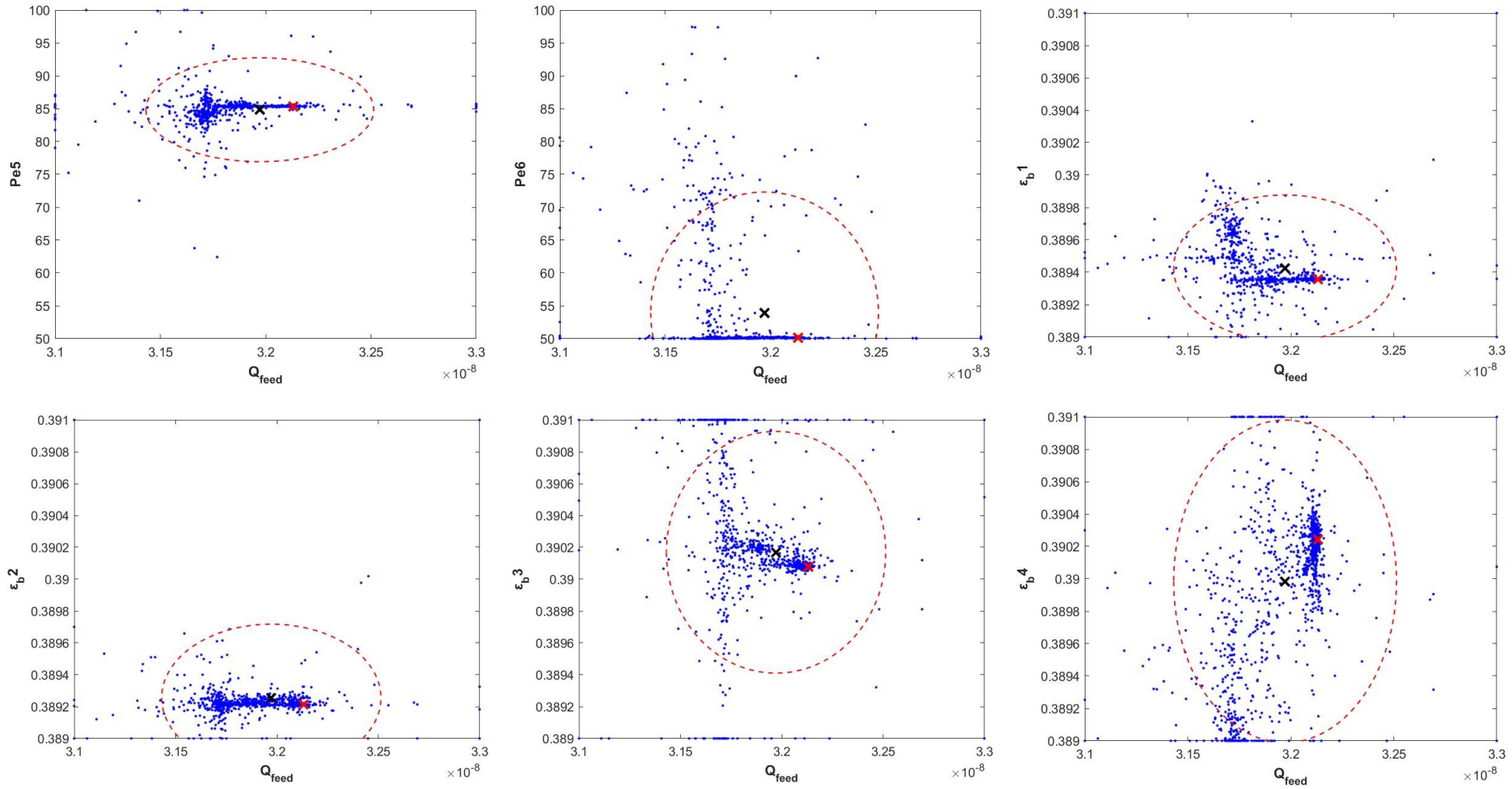

Figure B. 10 - Confidence Regions for each pair of estimable parameters – Pairs containing Q_{feed} (Continuation)

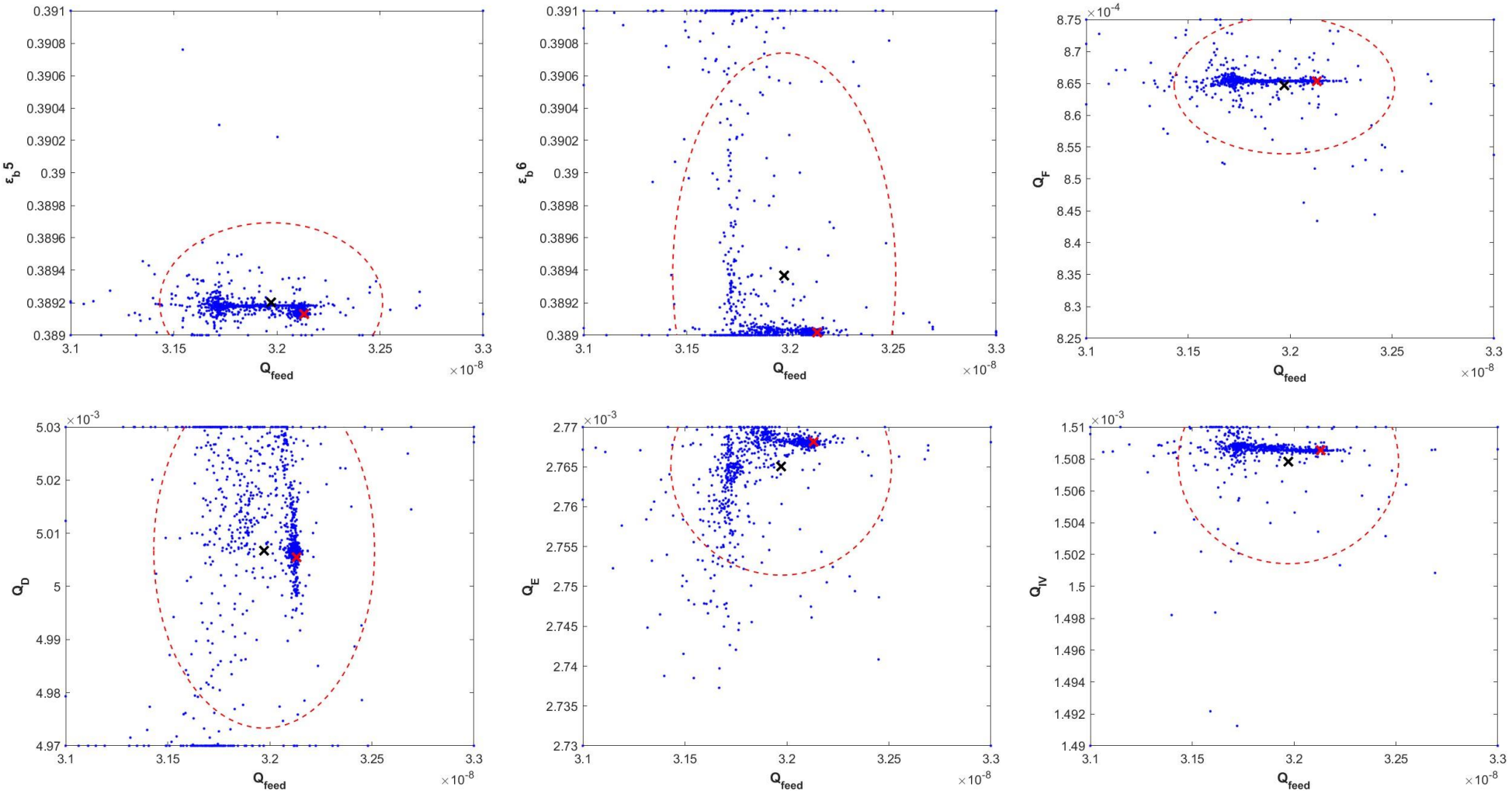

Figure B. 11 - Confidence Regions for each pair of estimable parameters – Pairs containing Q_{feed} (Continuation)

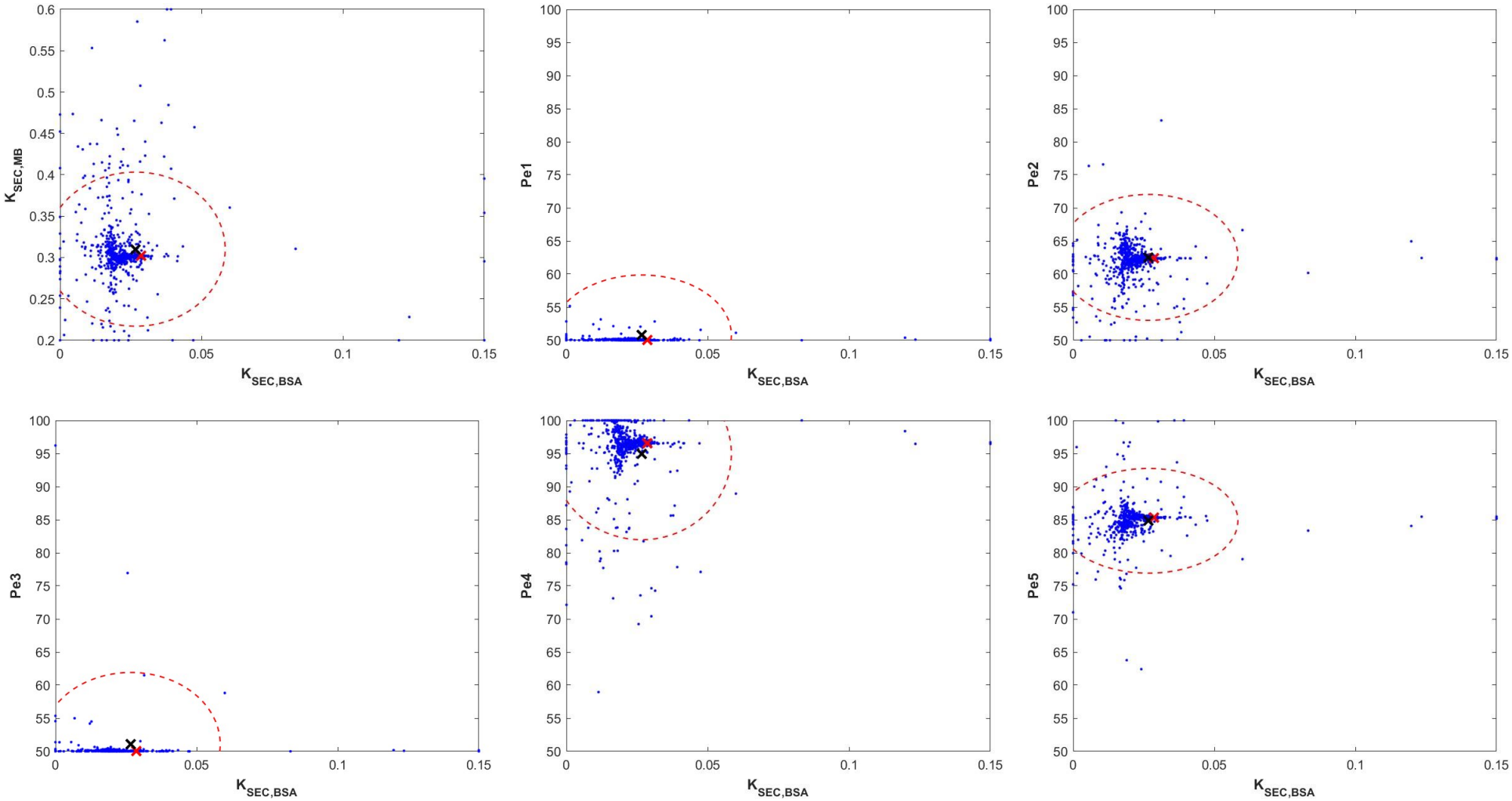

Figure B. 12 - Confidence Regions for each pair of estimable parameters – Pairs containing $K_{SEC,BSA}$

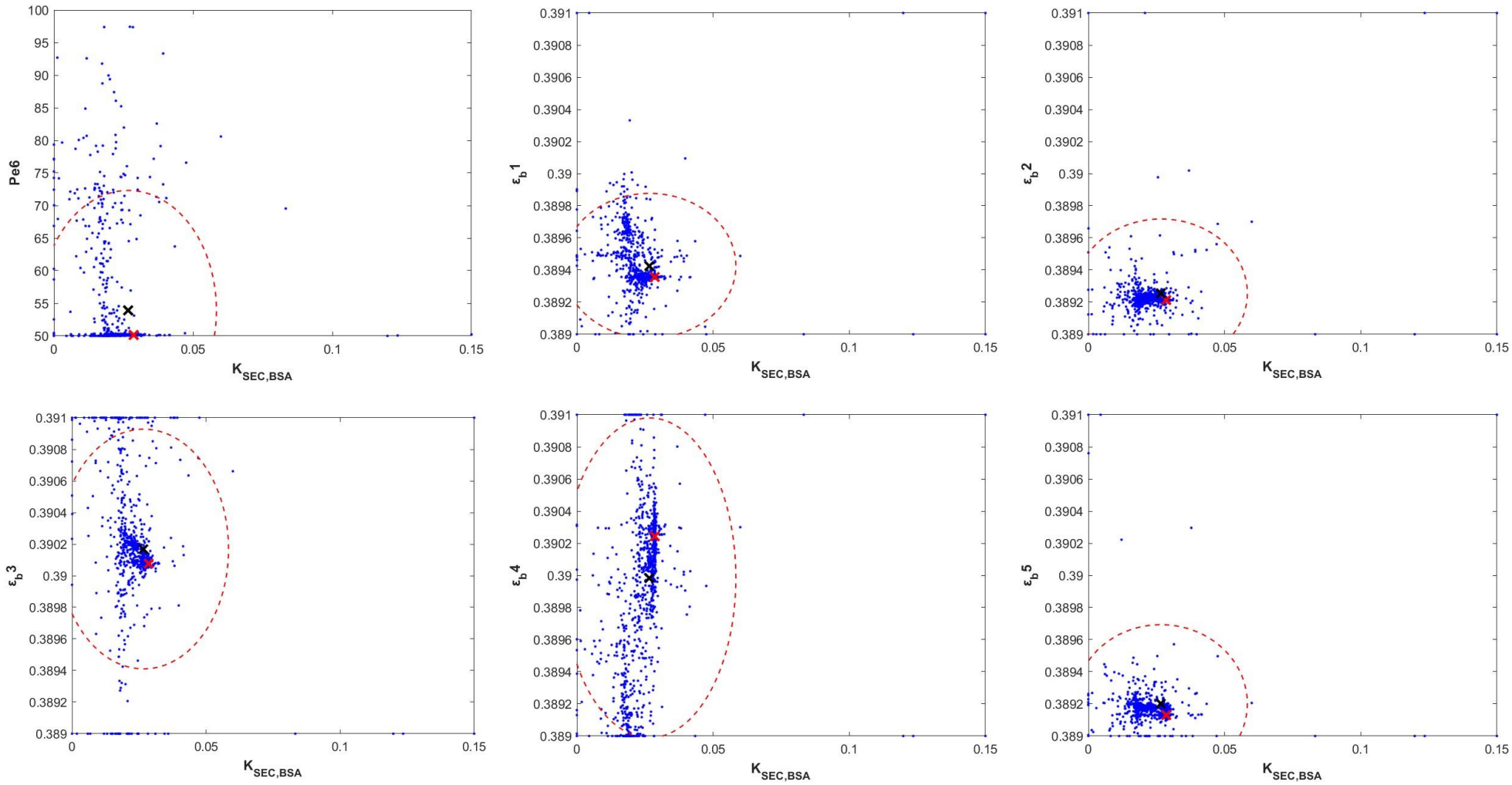

Figure B. 13 - Confidence Regions for each pair of estimable parameters – Pairs containing $K_{SEC,BSA}$ (Continuation)

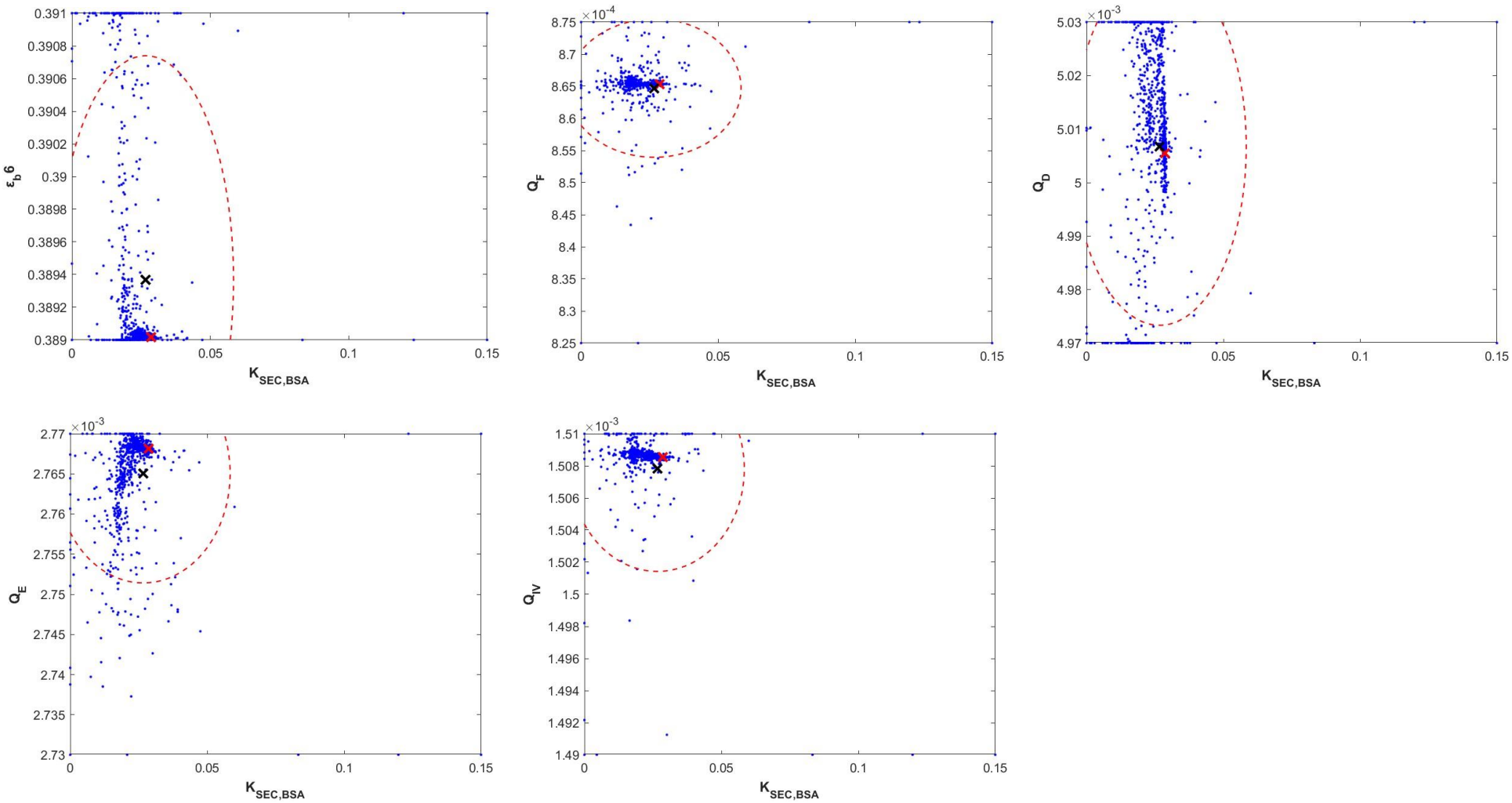

Figure B. 14 - Confidence Regions for each pair of estimable parameters – Pairs containing $K_{SEC,BSA}$ (Continuation)

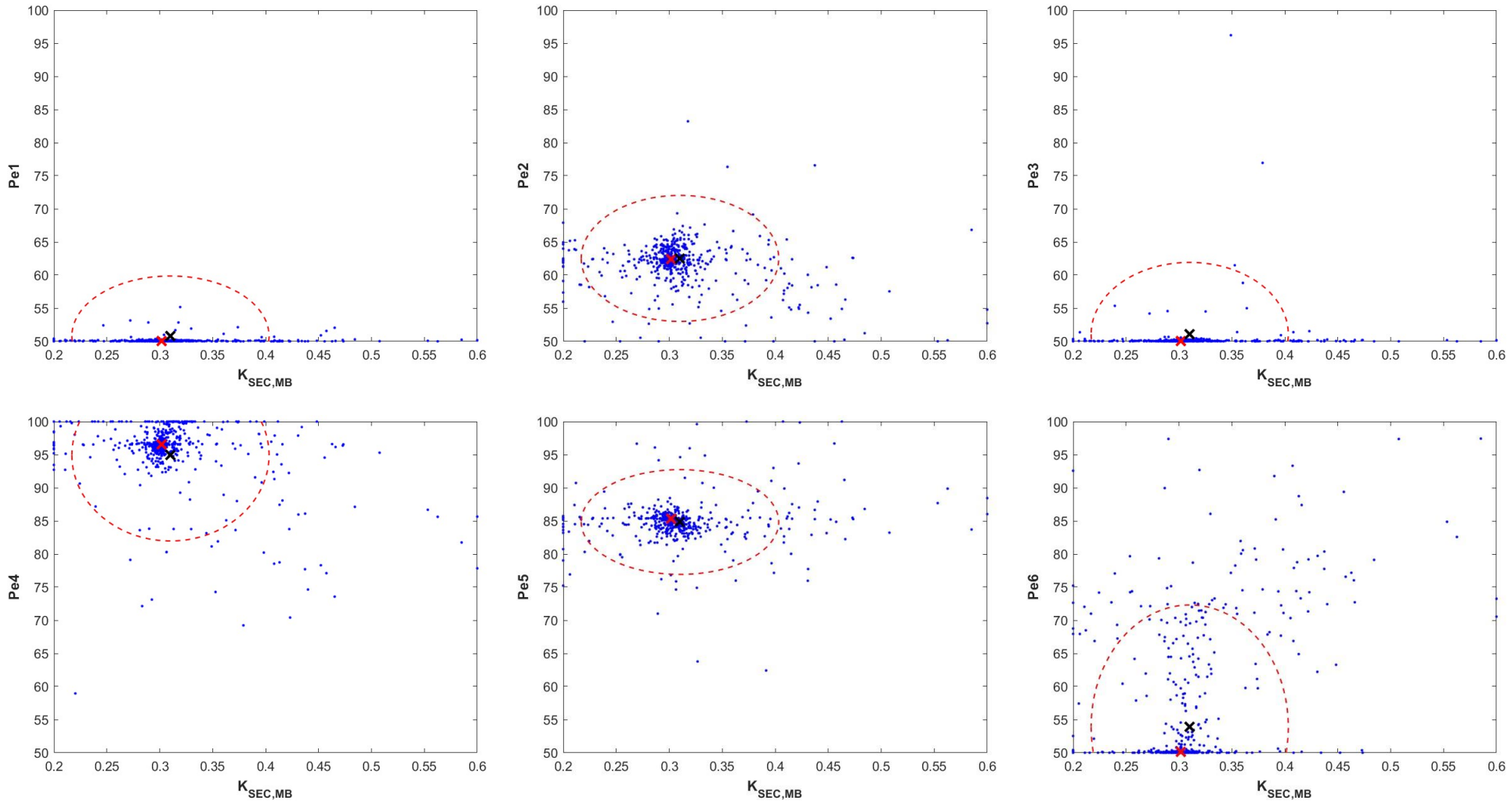

Figure B. 15 - Confidence Regions for each pair of estimable parameters – Pairs containing $K_{SEC,Mb}$

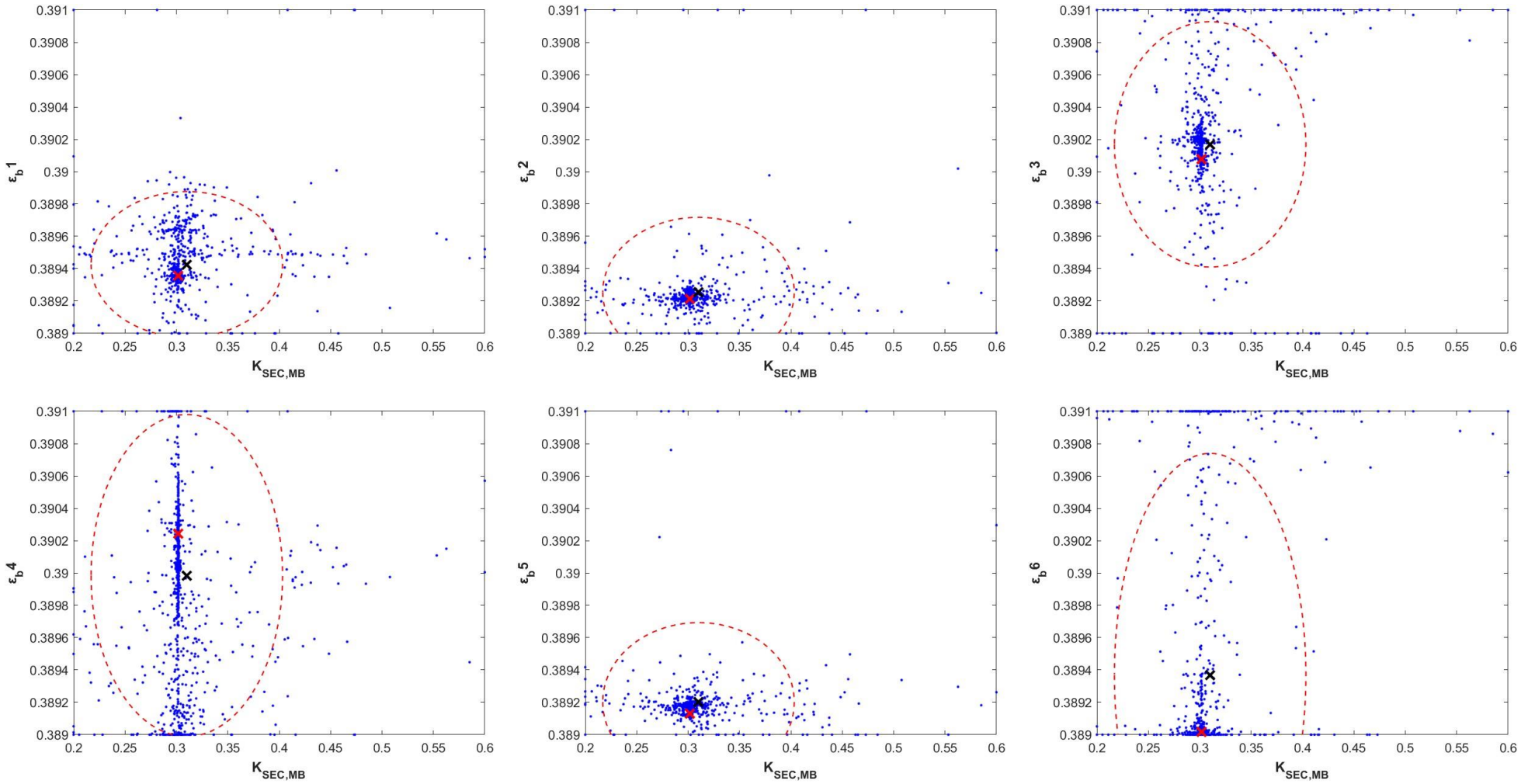

Figure B. 16 - Confidence Regions for each pair of estimable parameters – Pairs containing $K_{SEC,Mb}$ (Continuation)

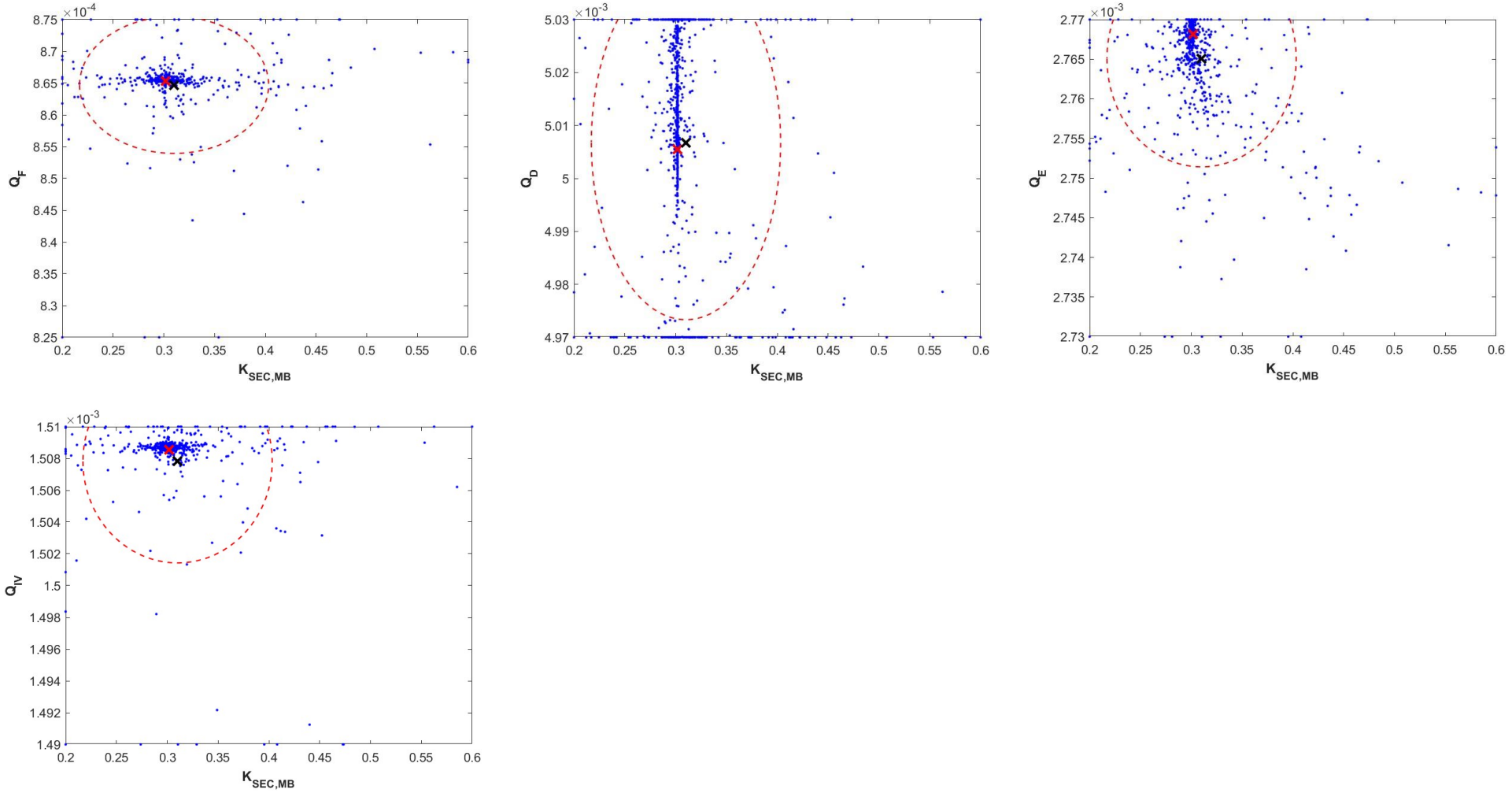

Figure B. 17 - Confidence Regions for each pair of estimable parameters – Pairs containing $K_{SEC,Mb}$ (Continuation)

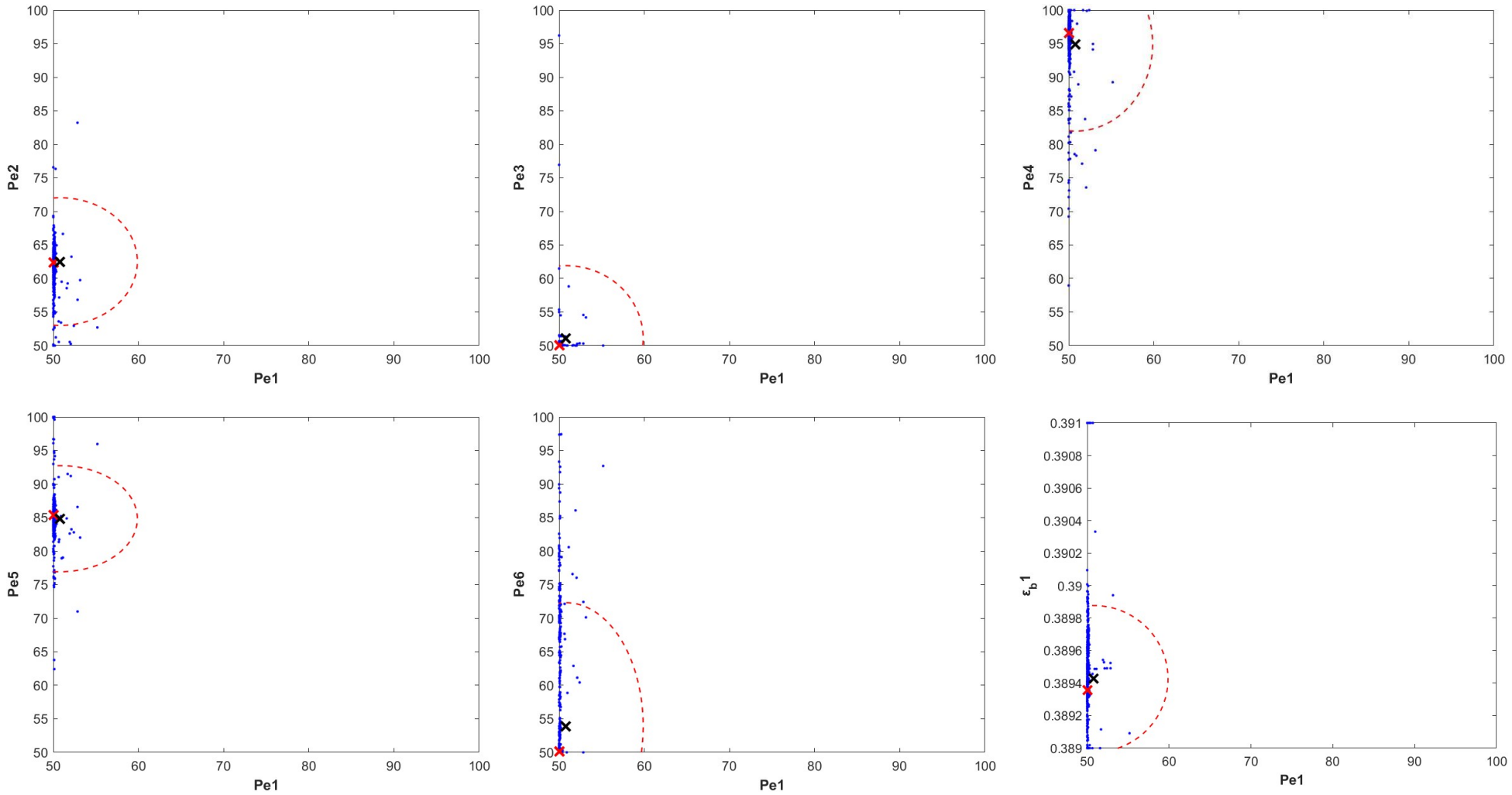

Figure B. 18 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_1

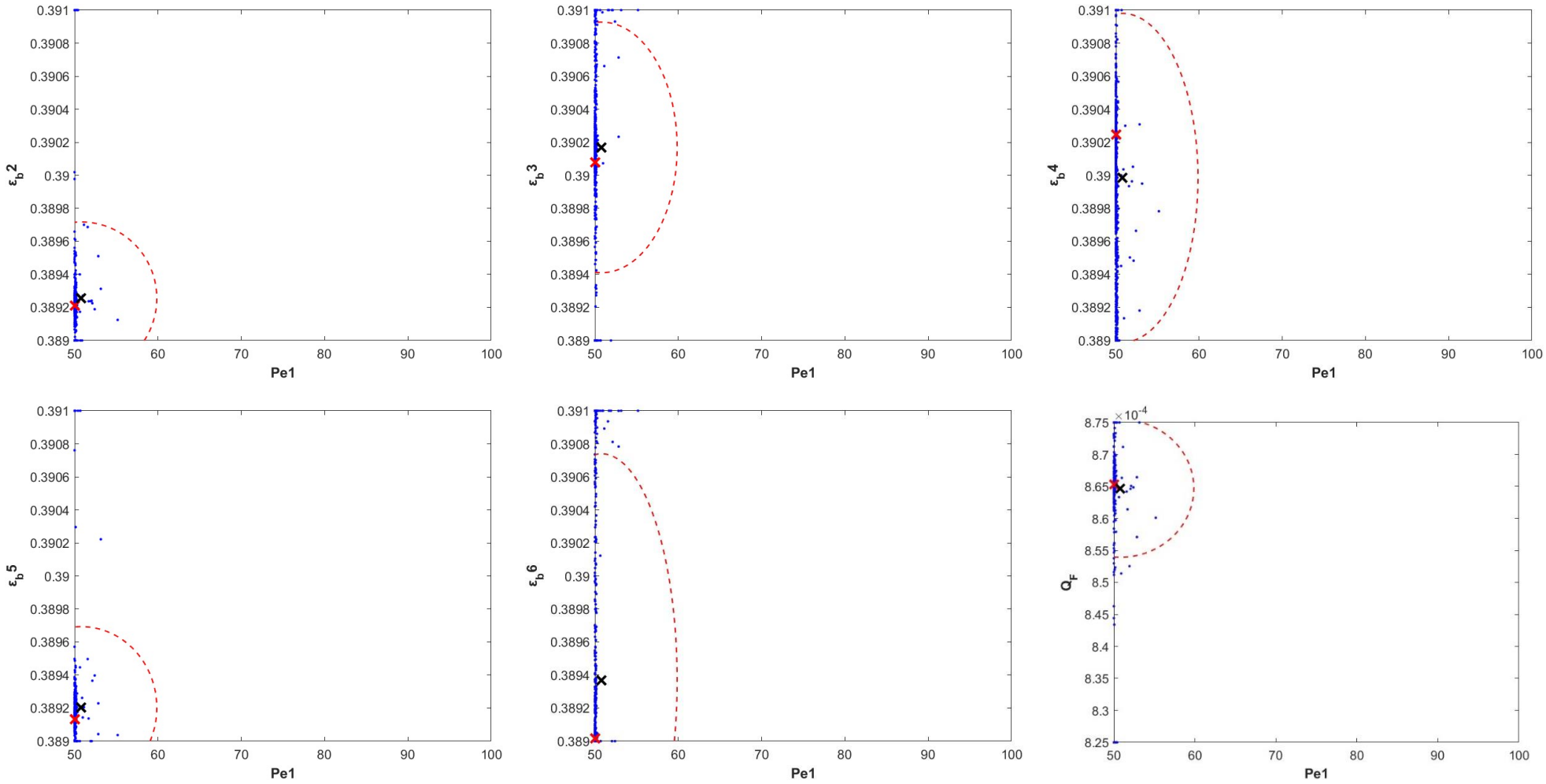

Figure B. 19 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_1 (Continuation)

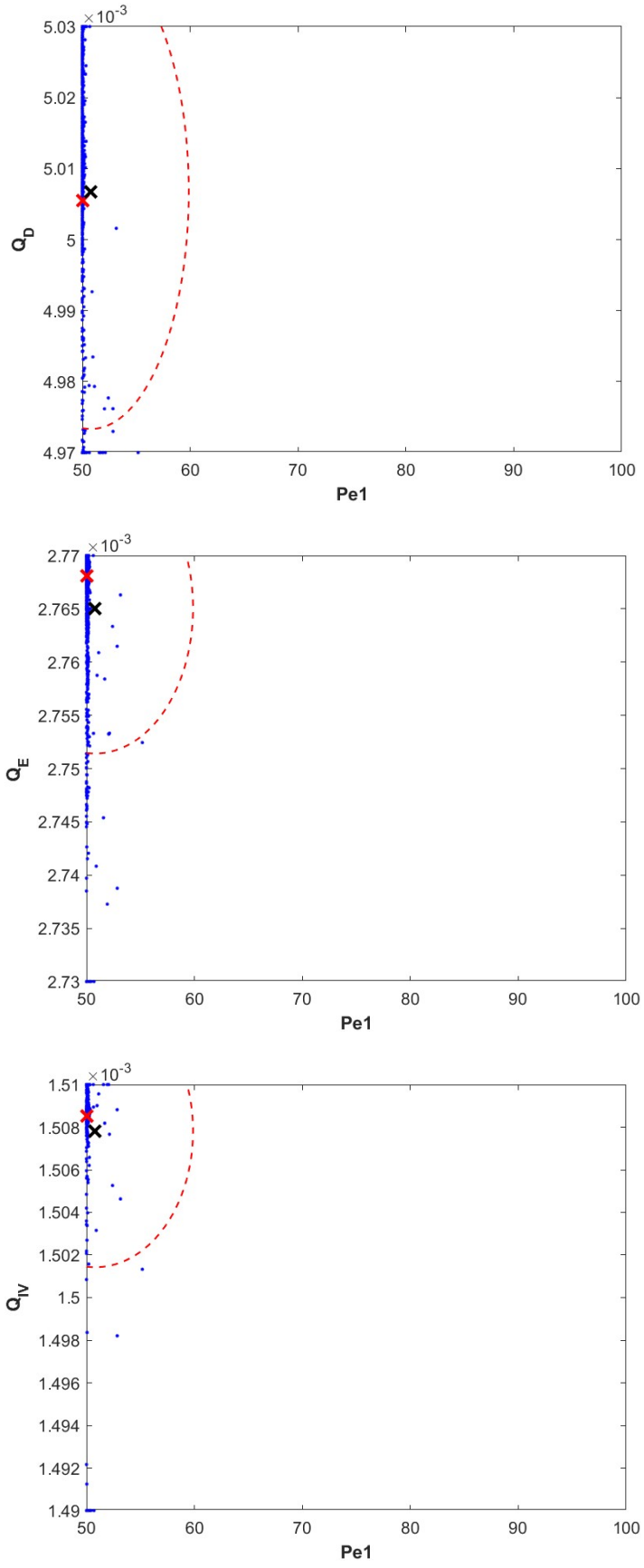

Figure B. 20 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_1 (Continuation)

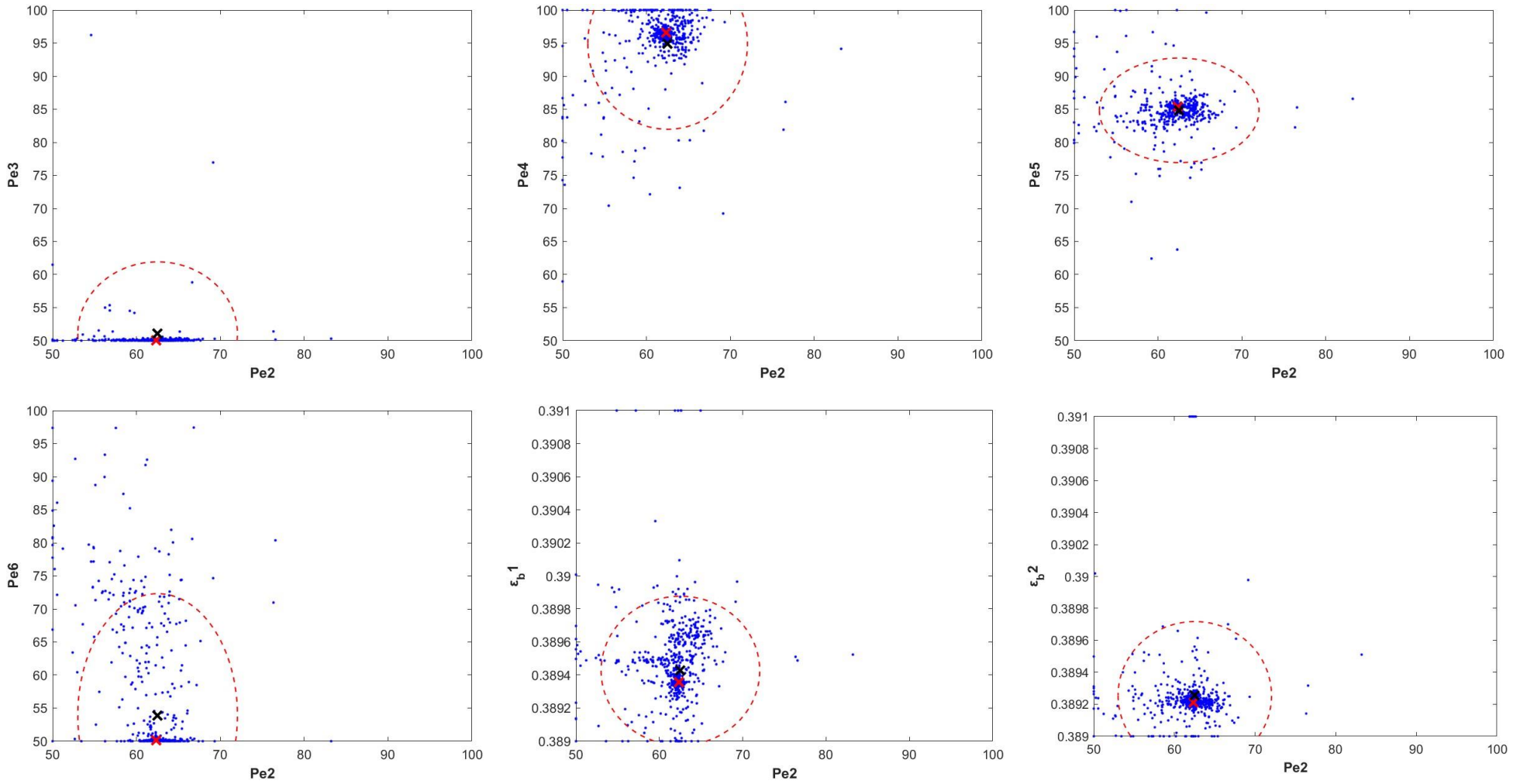

Figure B. 21 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_2

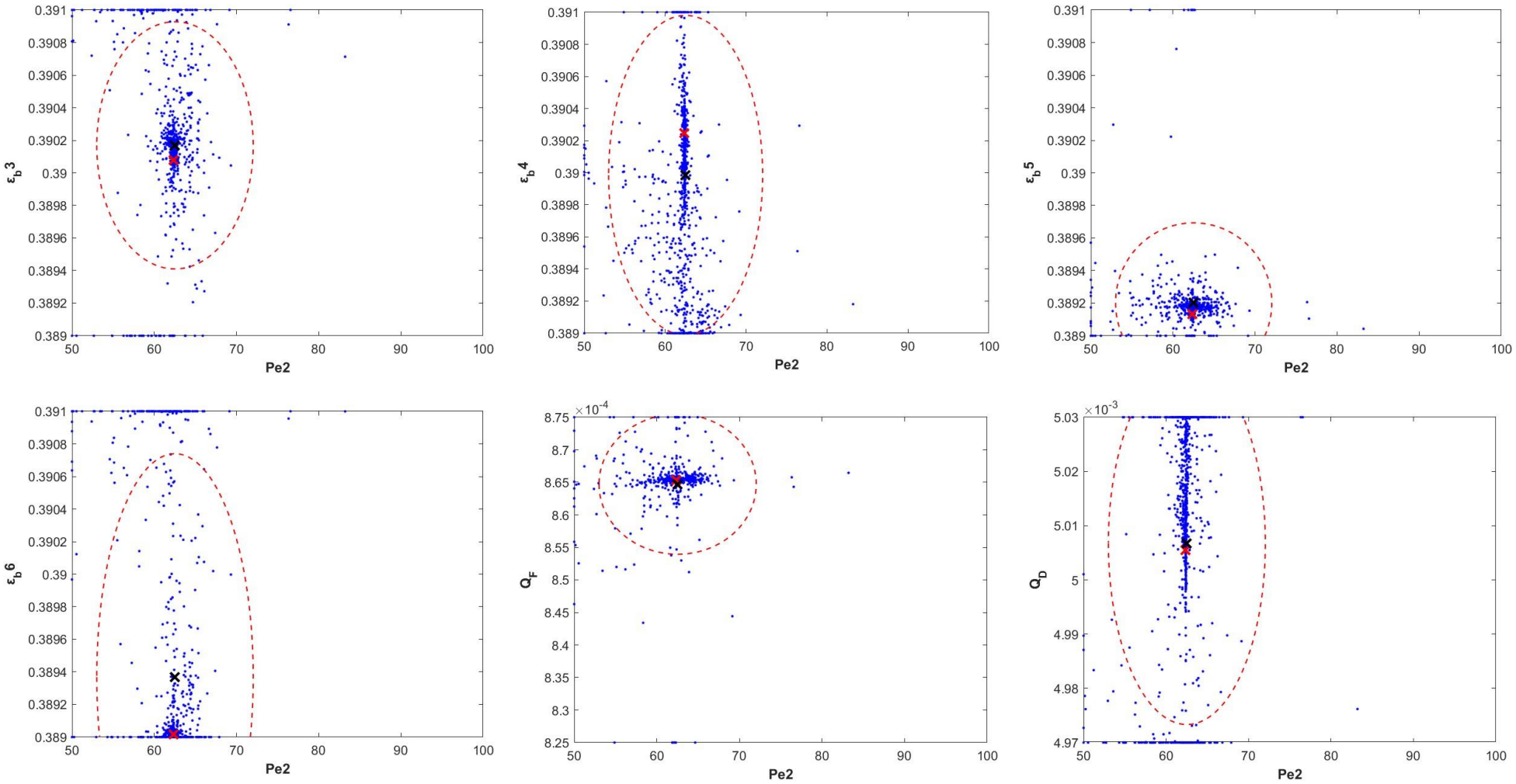

Figure B. 22 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_2 (Continuation)

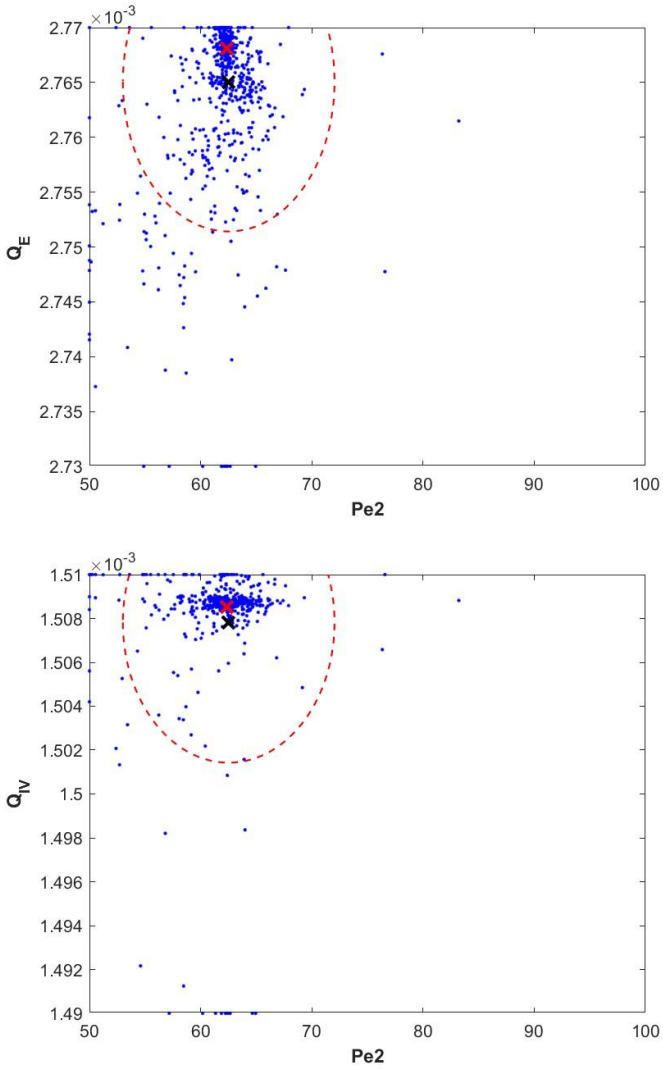

Figure B. 23 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_2 (Continuation)

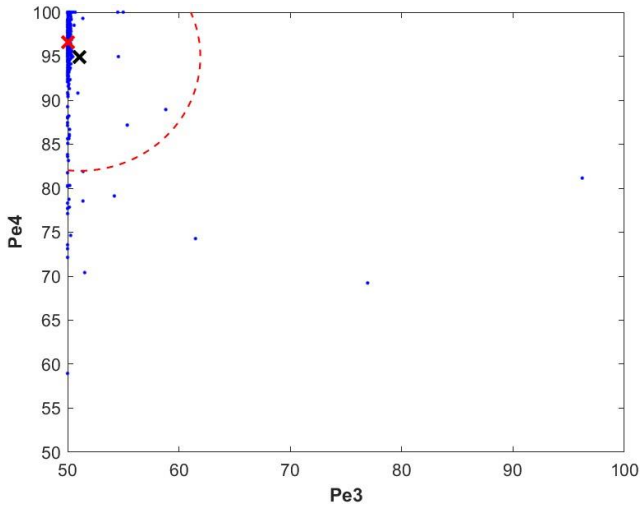

Figure B. 24 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_3

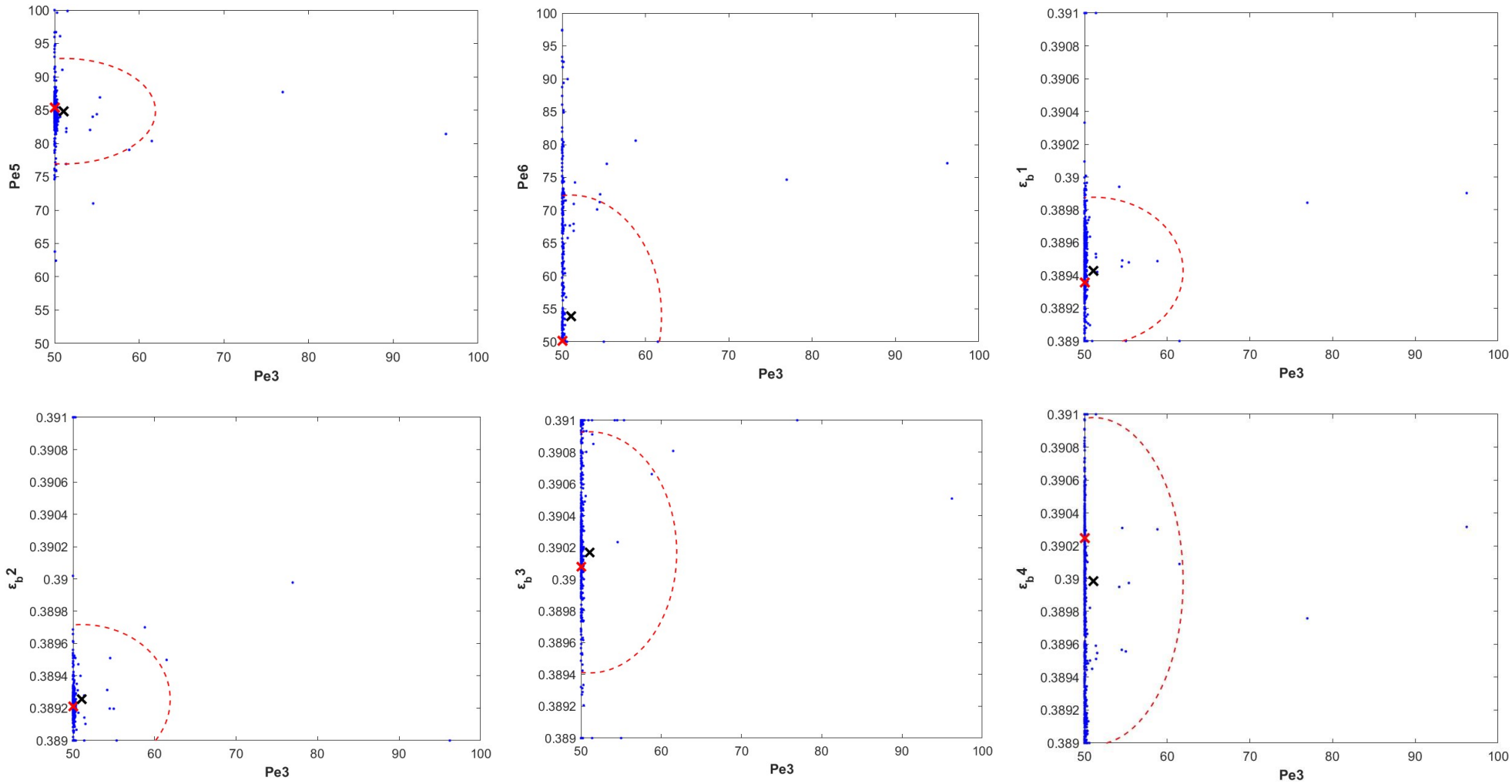

Figure B. 25 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_3 (Continuation)

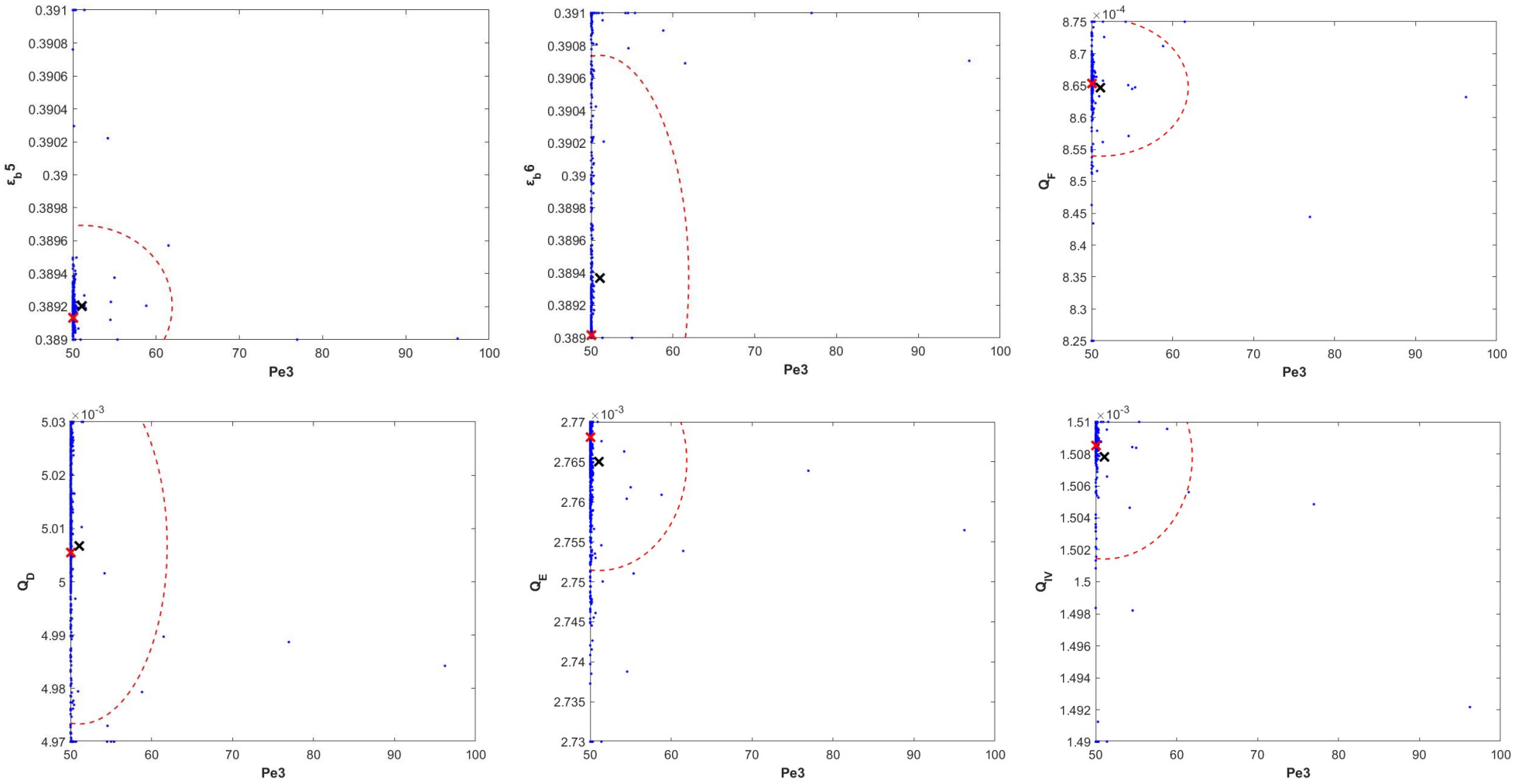

Figure B. 26 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_3 (Continuation)

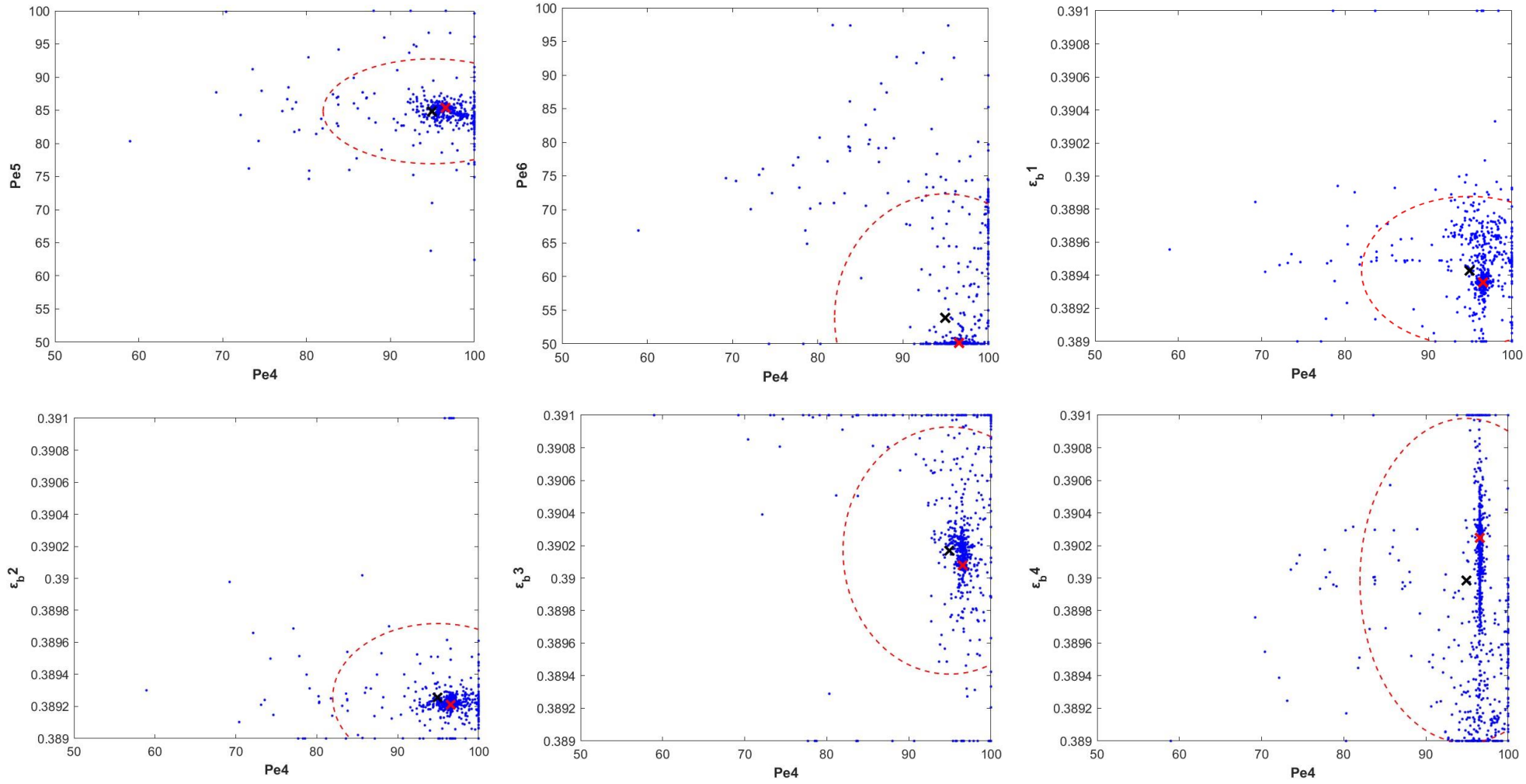

Figure B. 27 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_4

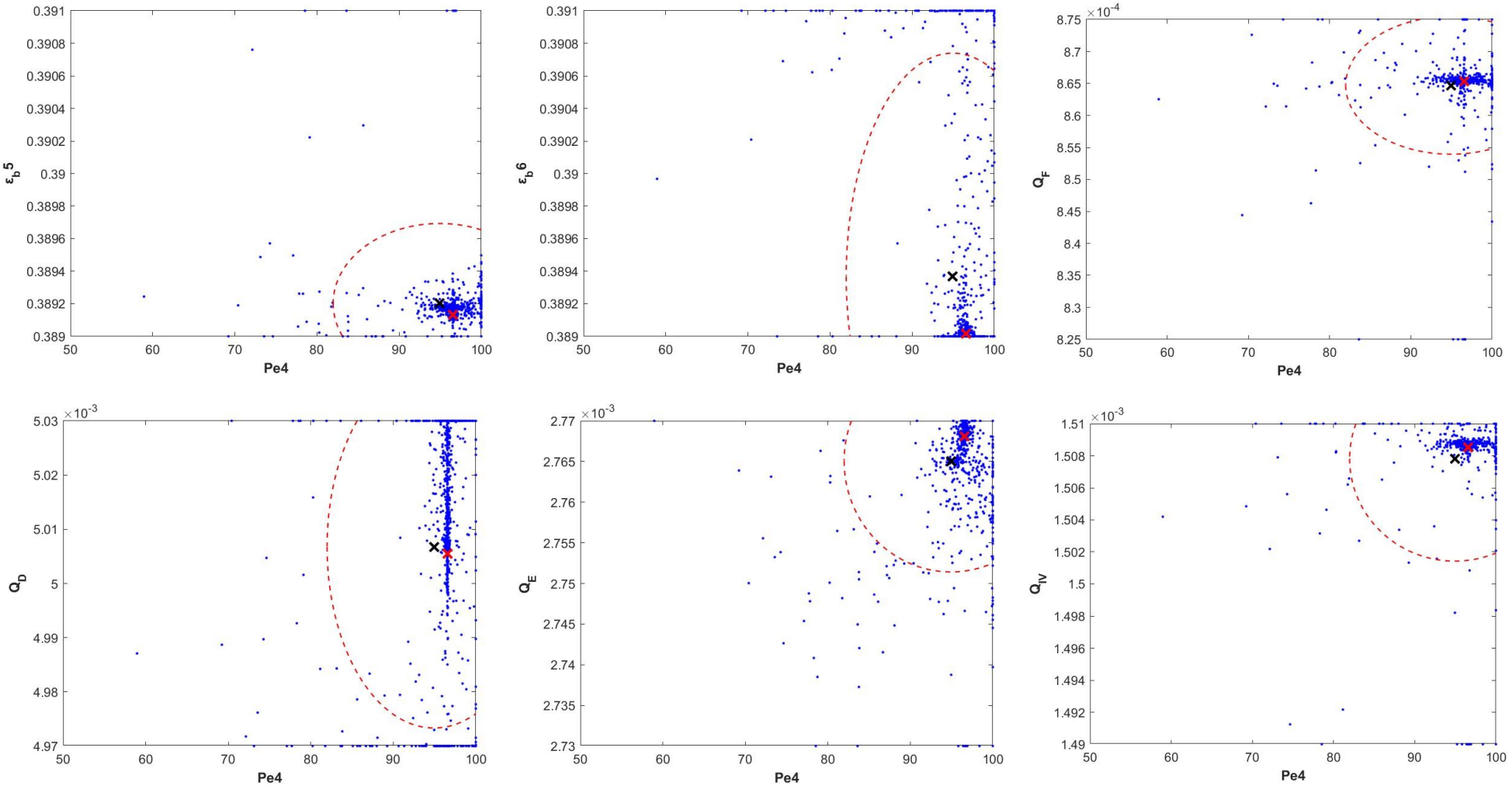

Figure B. 28 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_4 (Continuation)

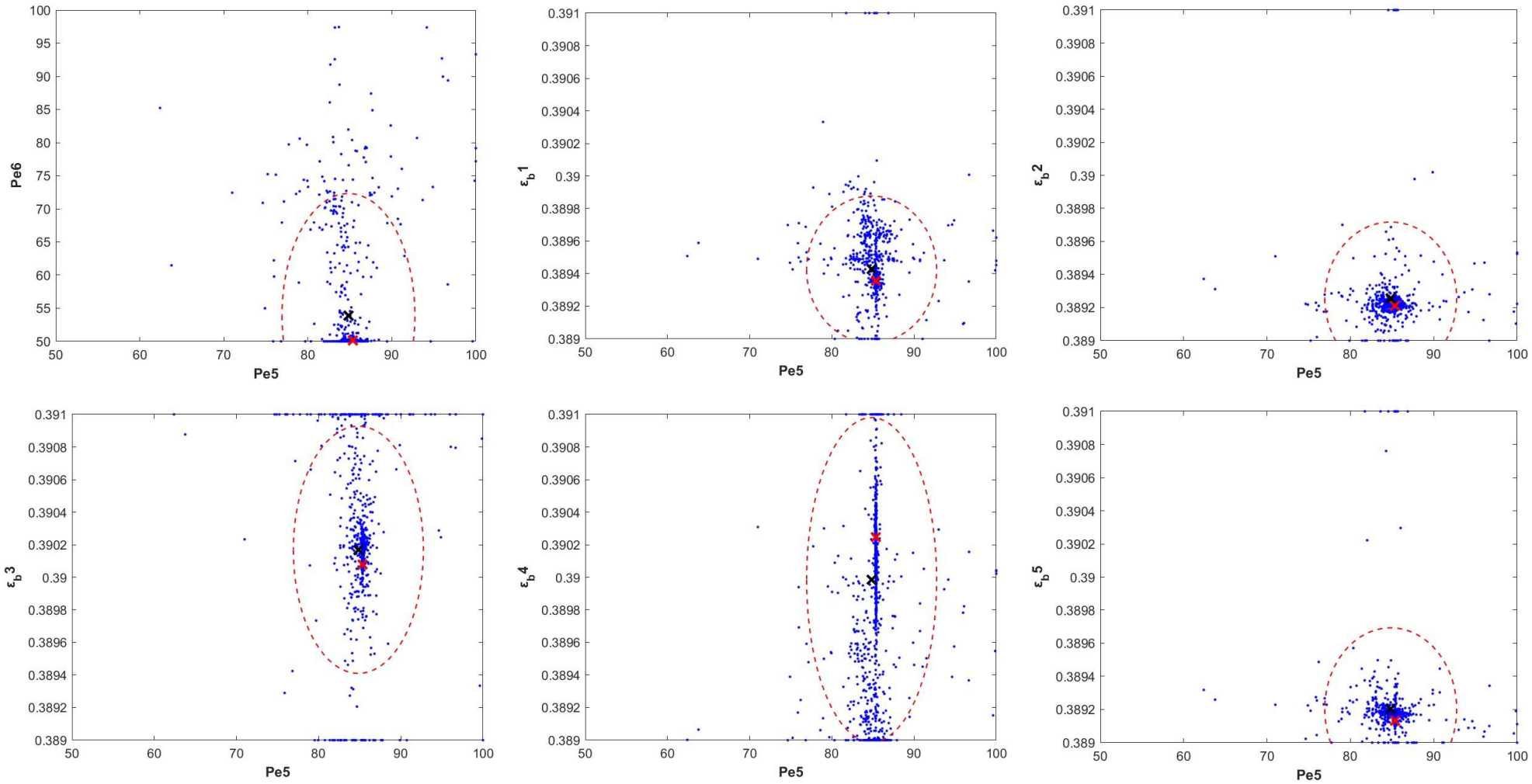

Figure B. 29 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_5

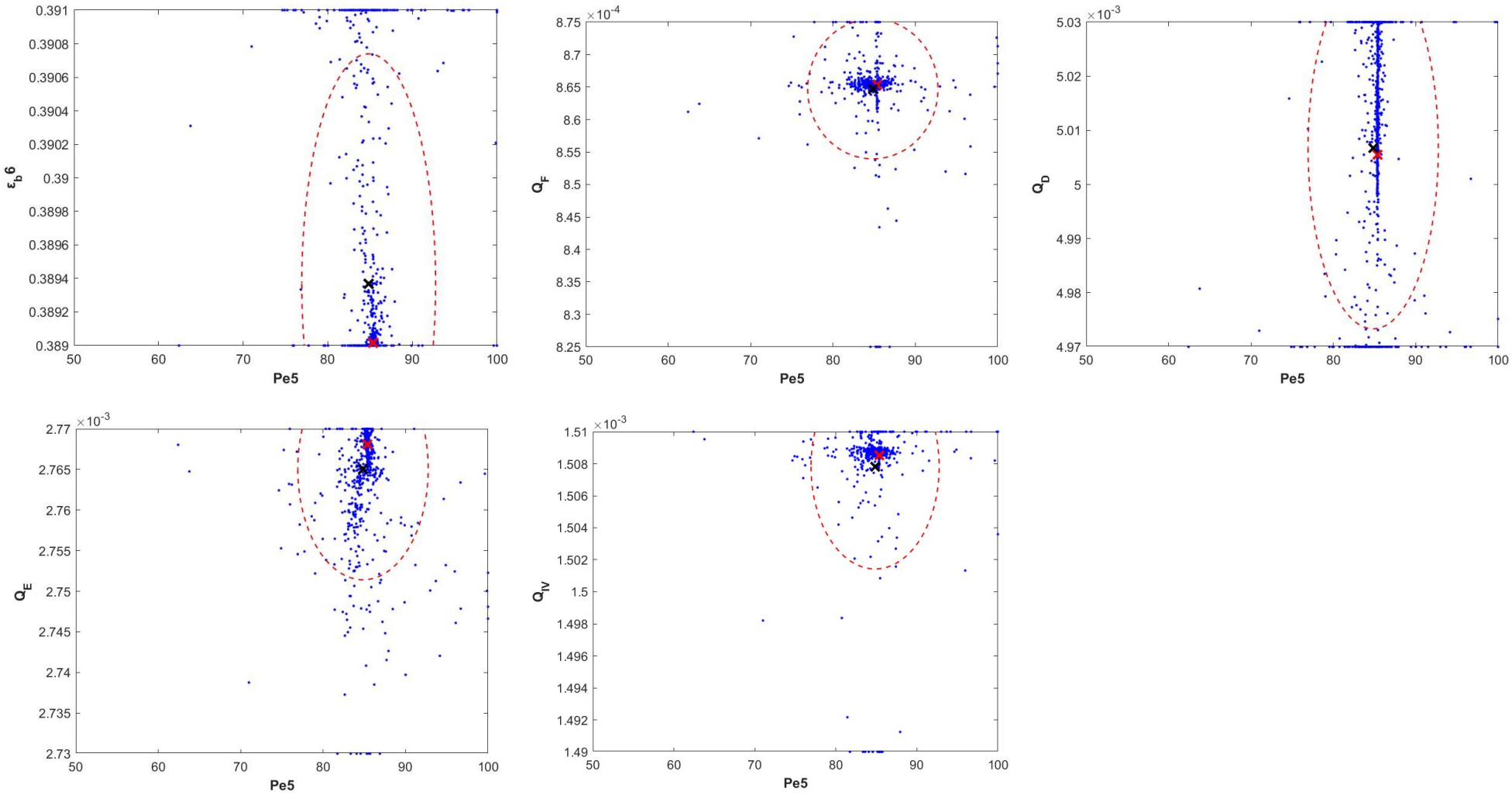

Figure B. 30 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_5 (Continuation)

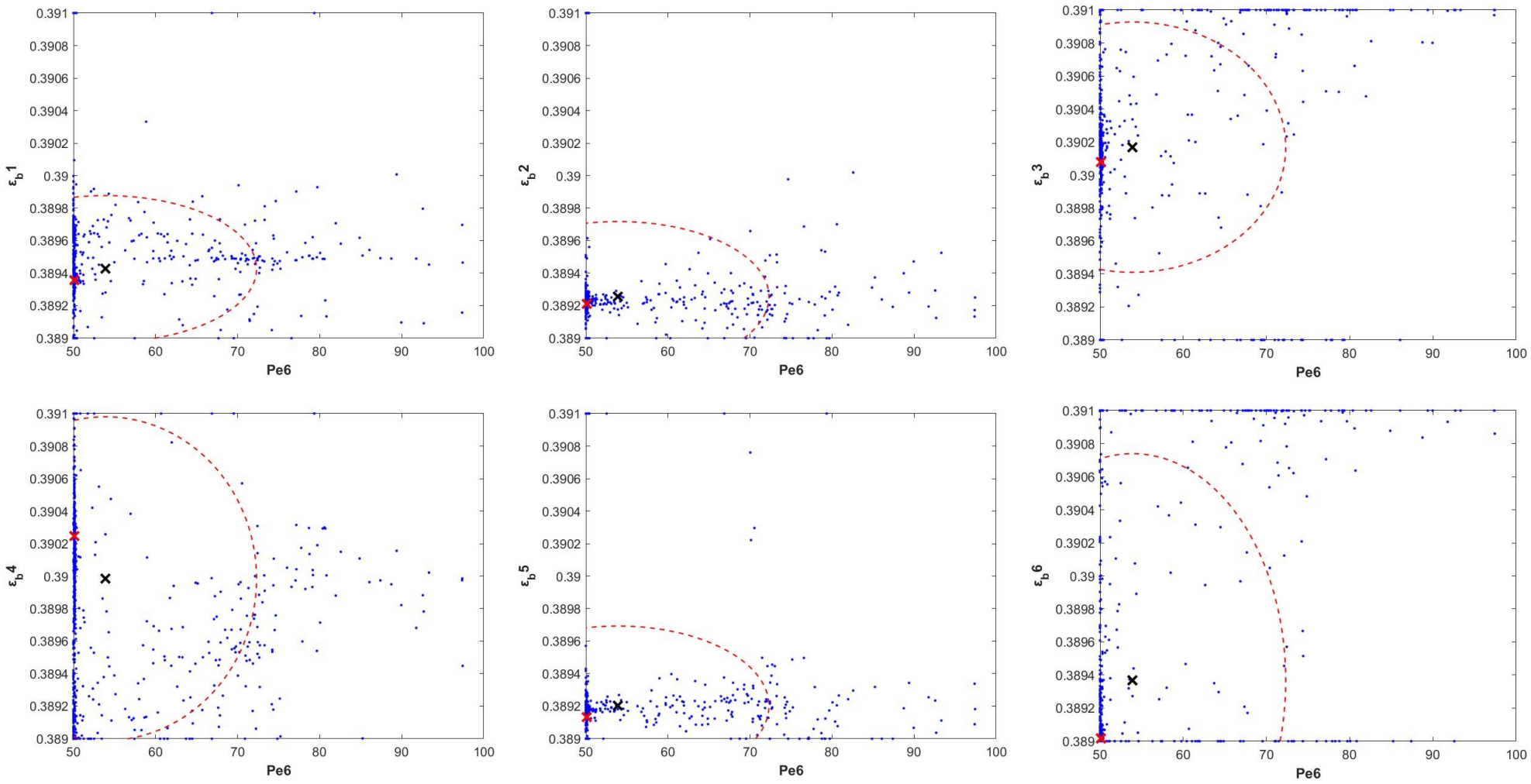

Figure B. 31 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_6

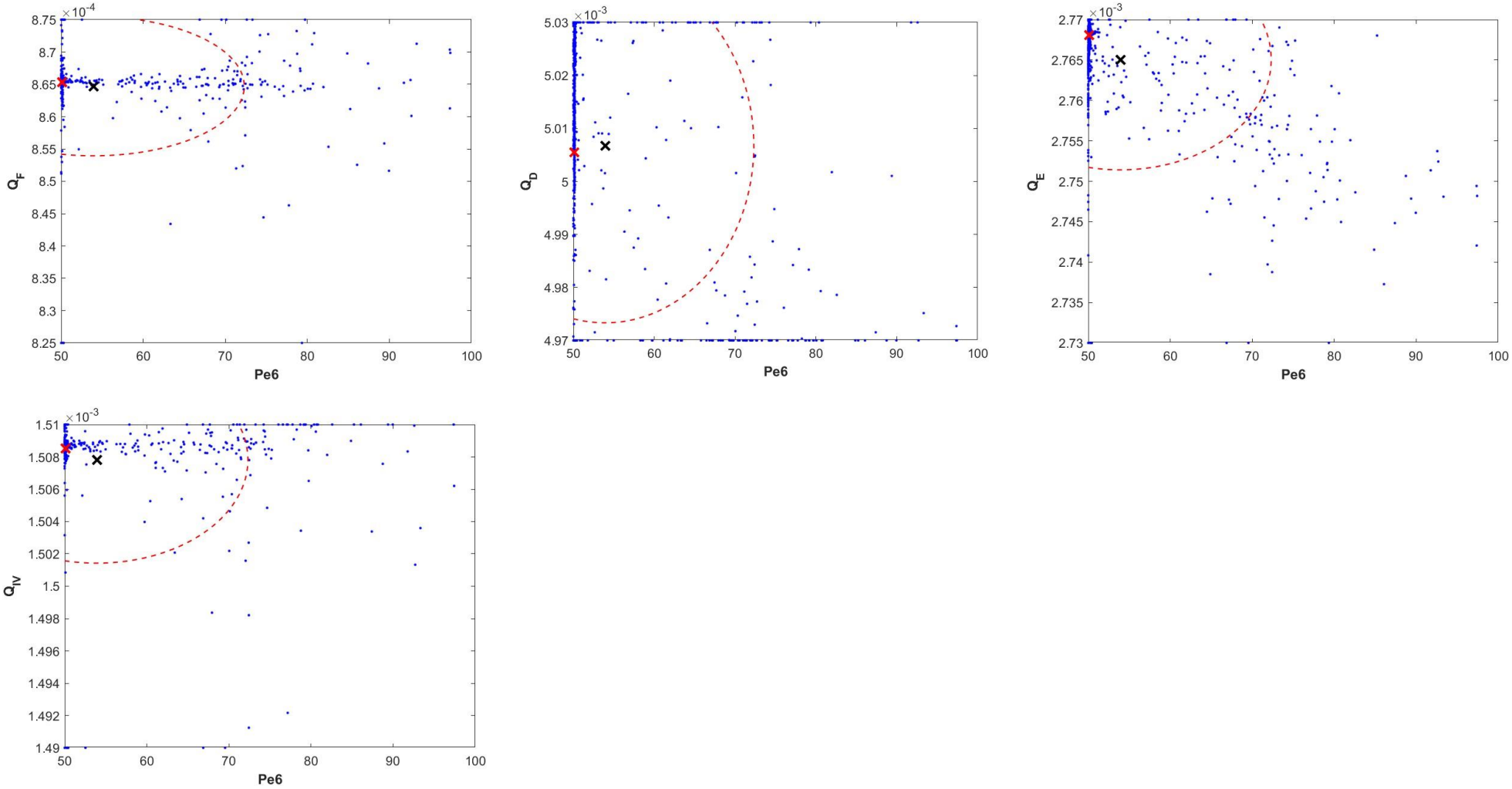

Figure B. 32 - Confidence Regions for each pair of estimable parameters – Pairs containing Pe_6 (Continuation)

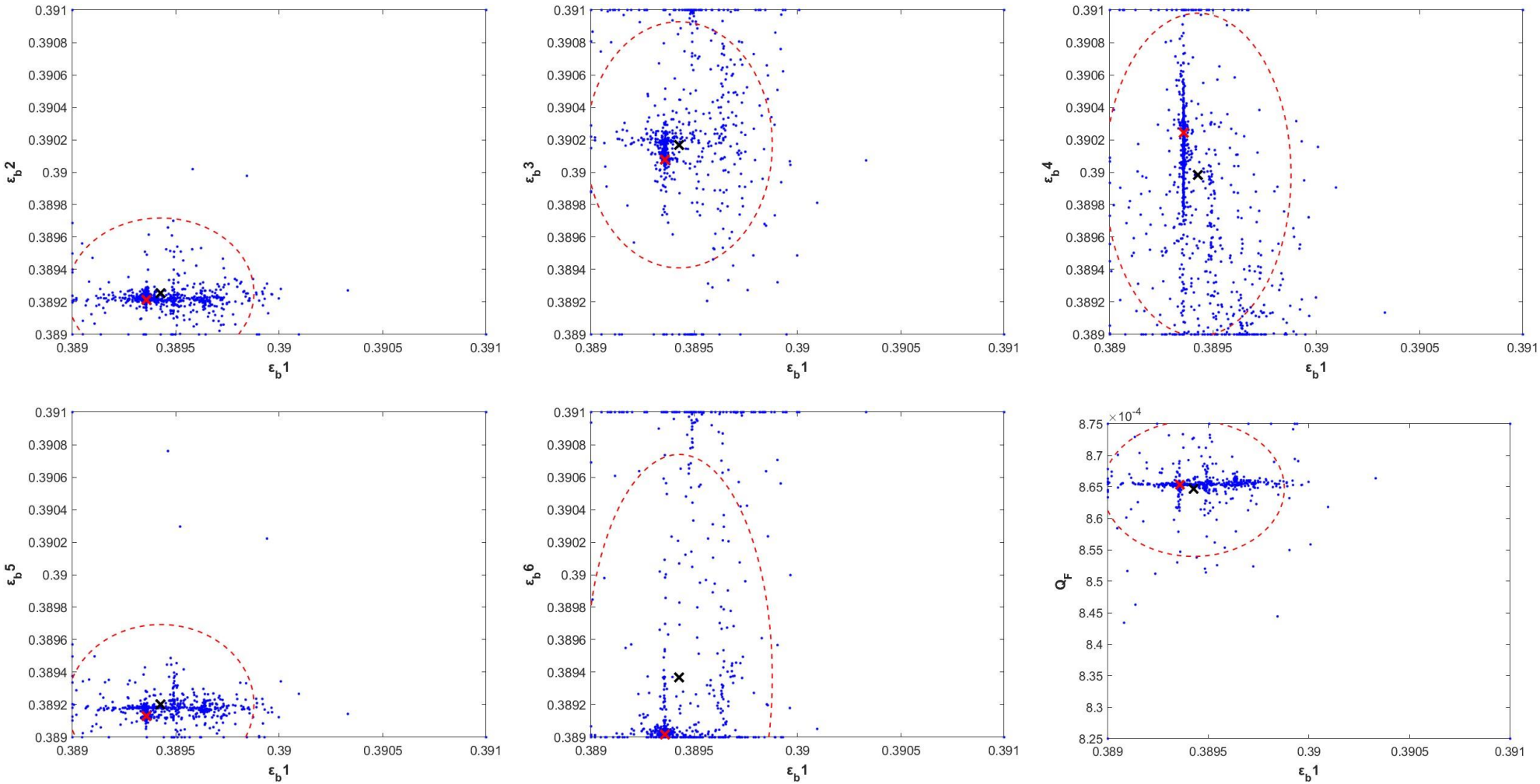

Figure B. 33 - Confidence Regions for each pair of estimable parameters – Pairs containing $\varepsilon_{b,1}$

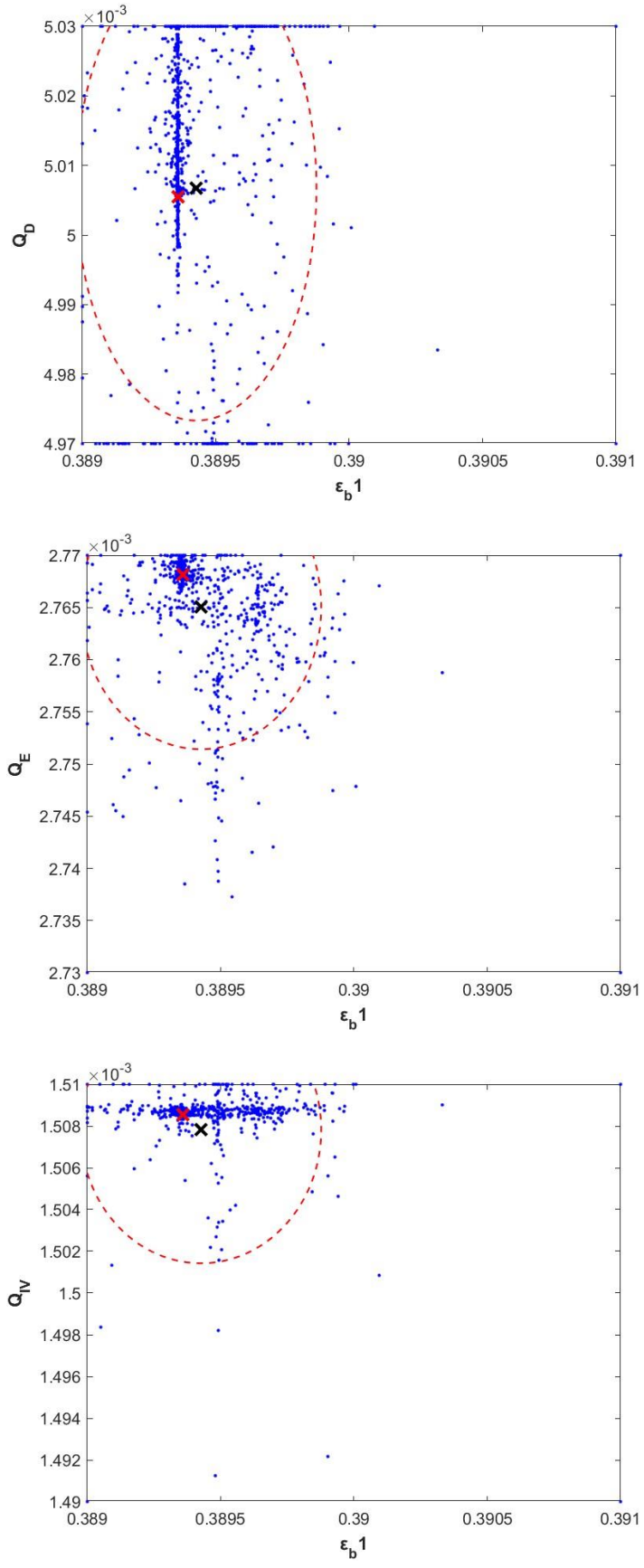

Figure B. 34 - Confidence Regions for each pair of estimable parameters – Pairs containing $\varepsilon_{b,1}$ (Continuation)

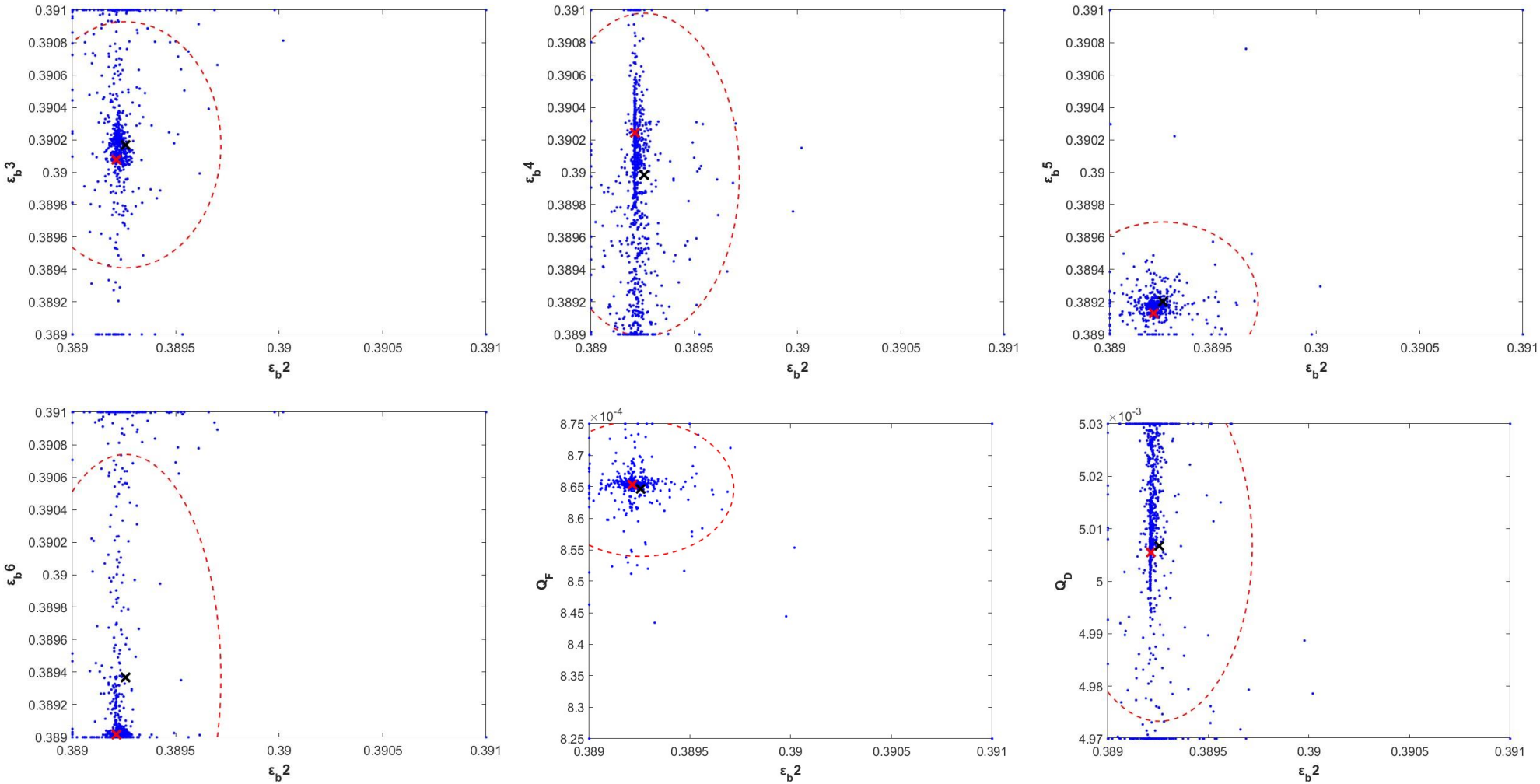

Figure B. 35 - Confidence Regions for each pair of estimable parameters – Pairs containing $\varepsilon_{b,2}$

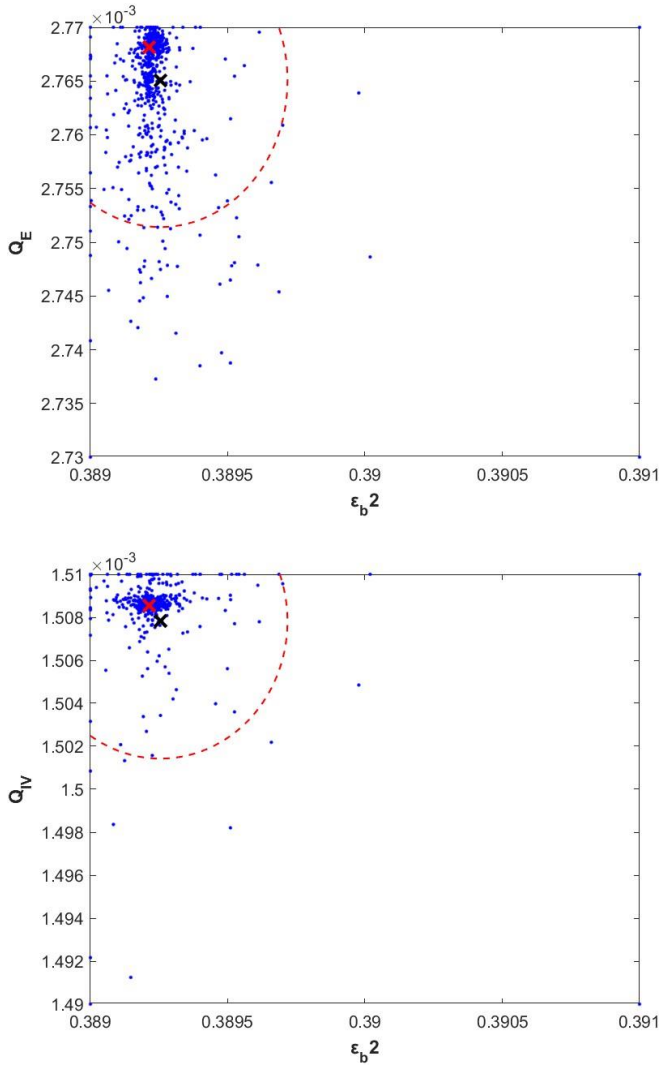

Figure B. 36 - Confidence Regions for each pair of estimable parameters – Pairs containing $\varepsilon_{b,2}$ (Continuation)

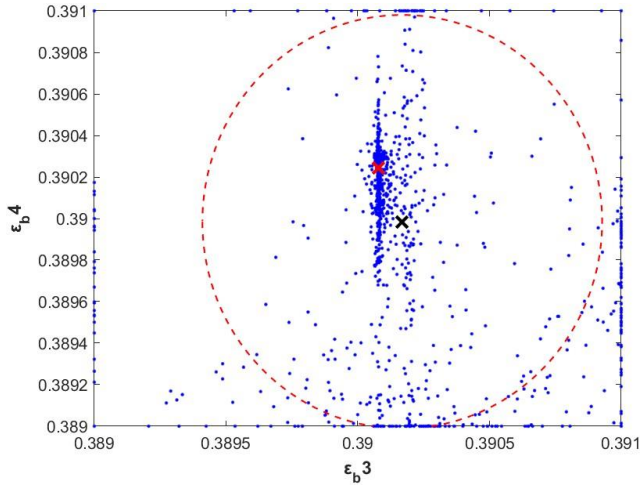

Figure B. 37 - Confidence Regions for each pair of estimable parameters – Pairs containing $\varepsilon_{b,3}$

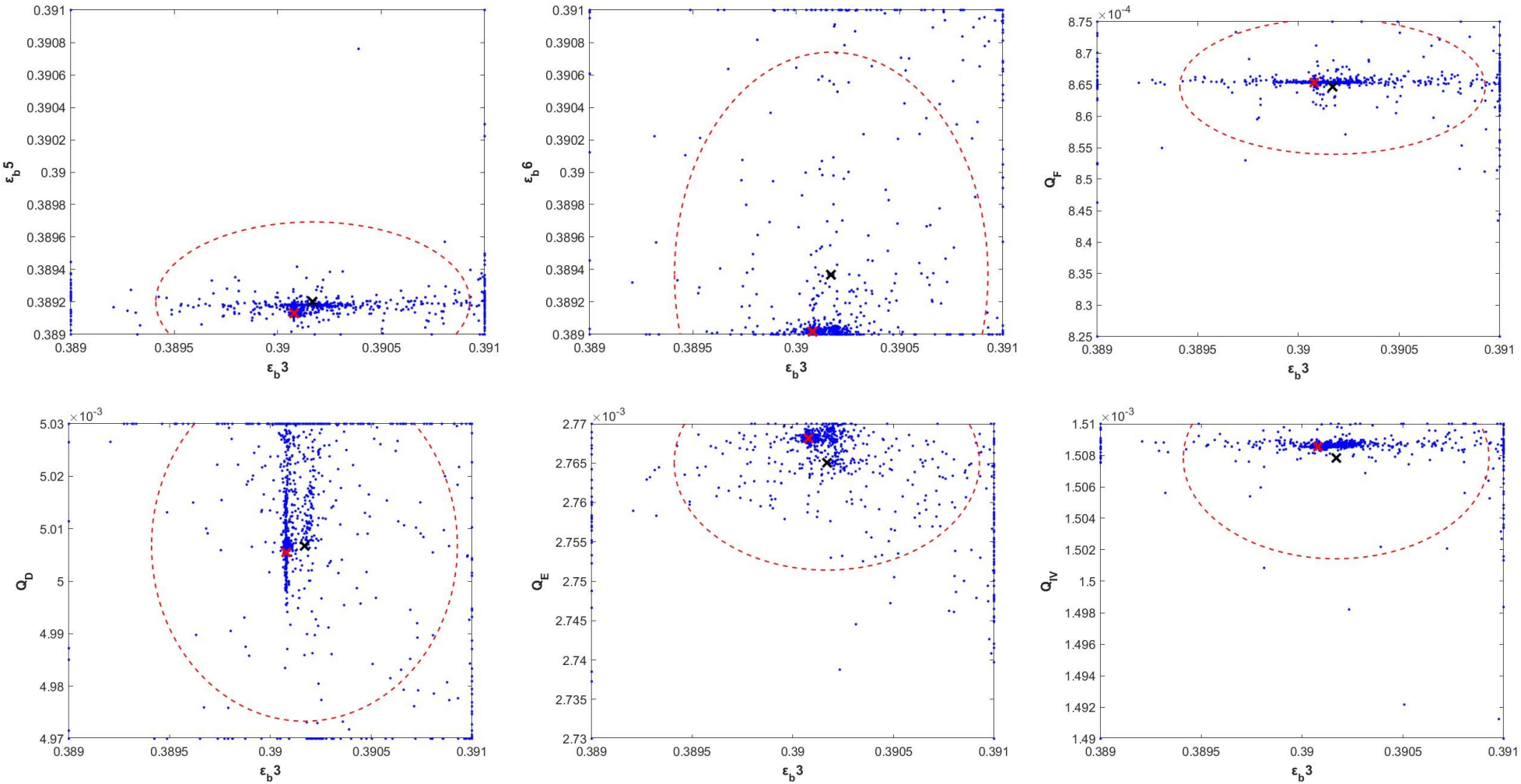

Figure B. 38 - Confidence Regions for each pair of estimable parameters – Pairs containing $\varepsilon_{b,3}$ (Continuation)

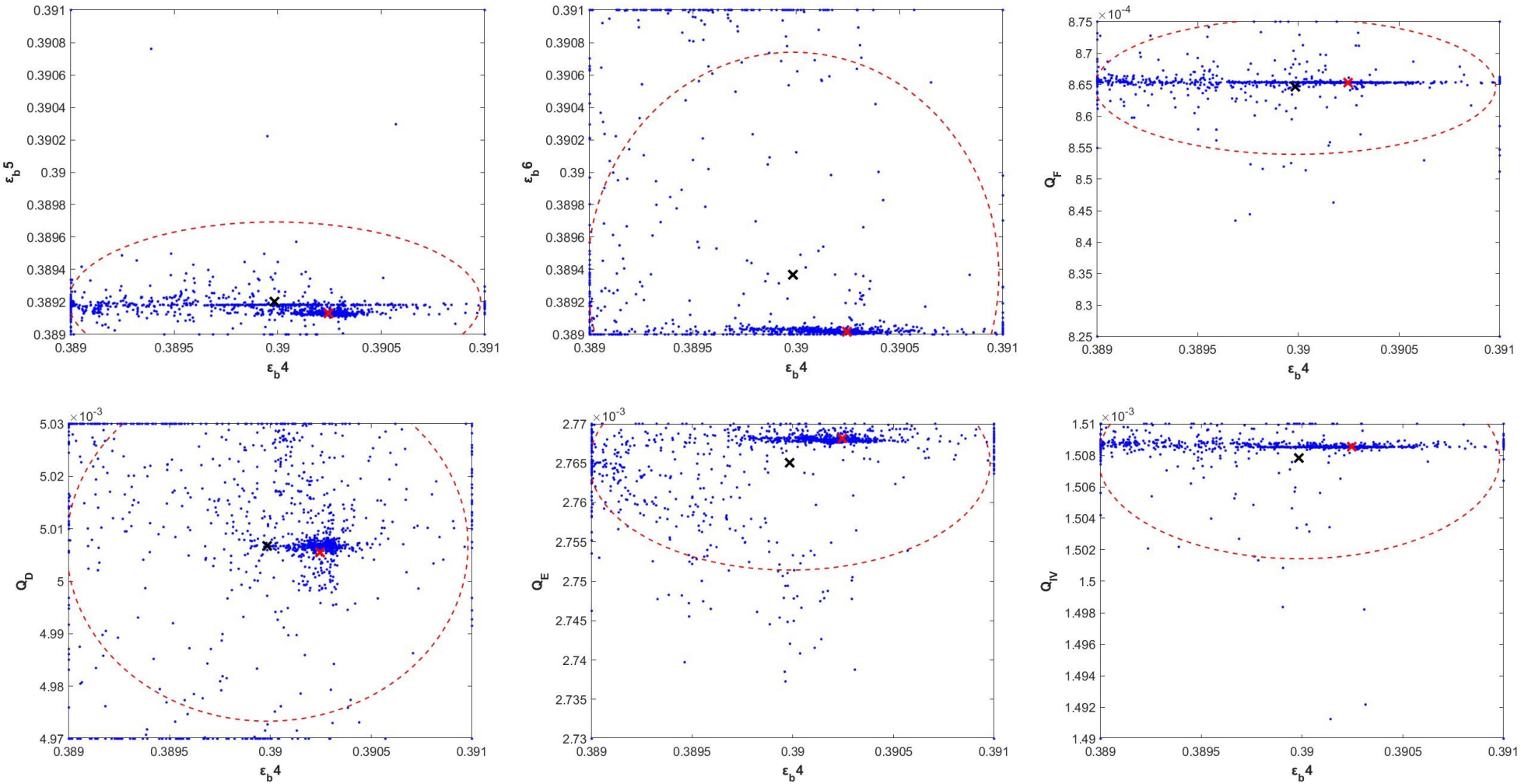

Figure B. 39 - Confidence Regions for each pair of estimable parameters – Pairs containing $\varepsilon_{b,4}$

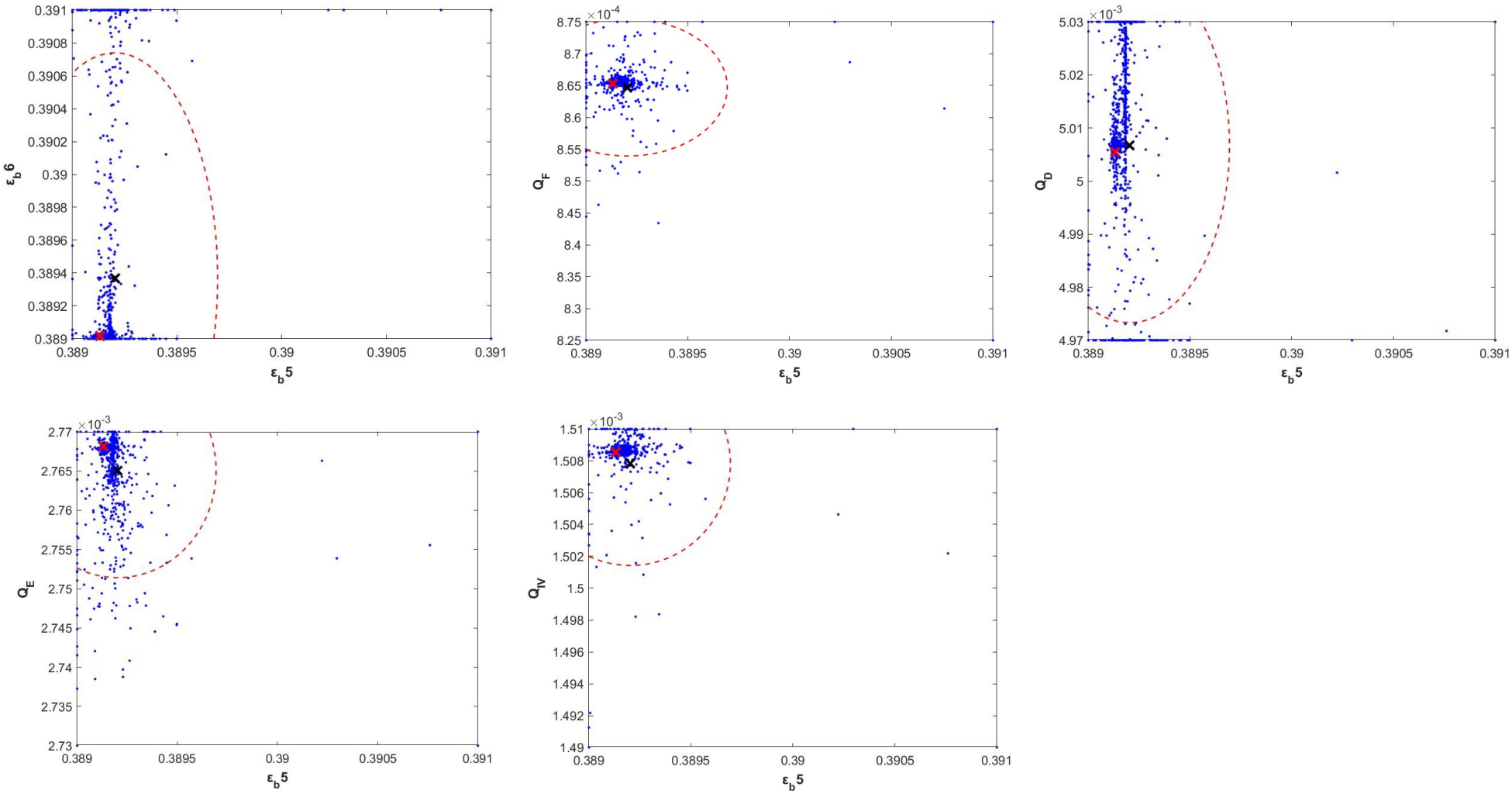

Figure B. 40 - Confidence Regions for each pair of estimable parameters – Pairs containing $\varepsilon_{b,5}$

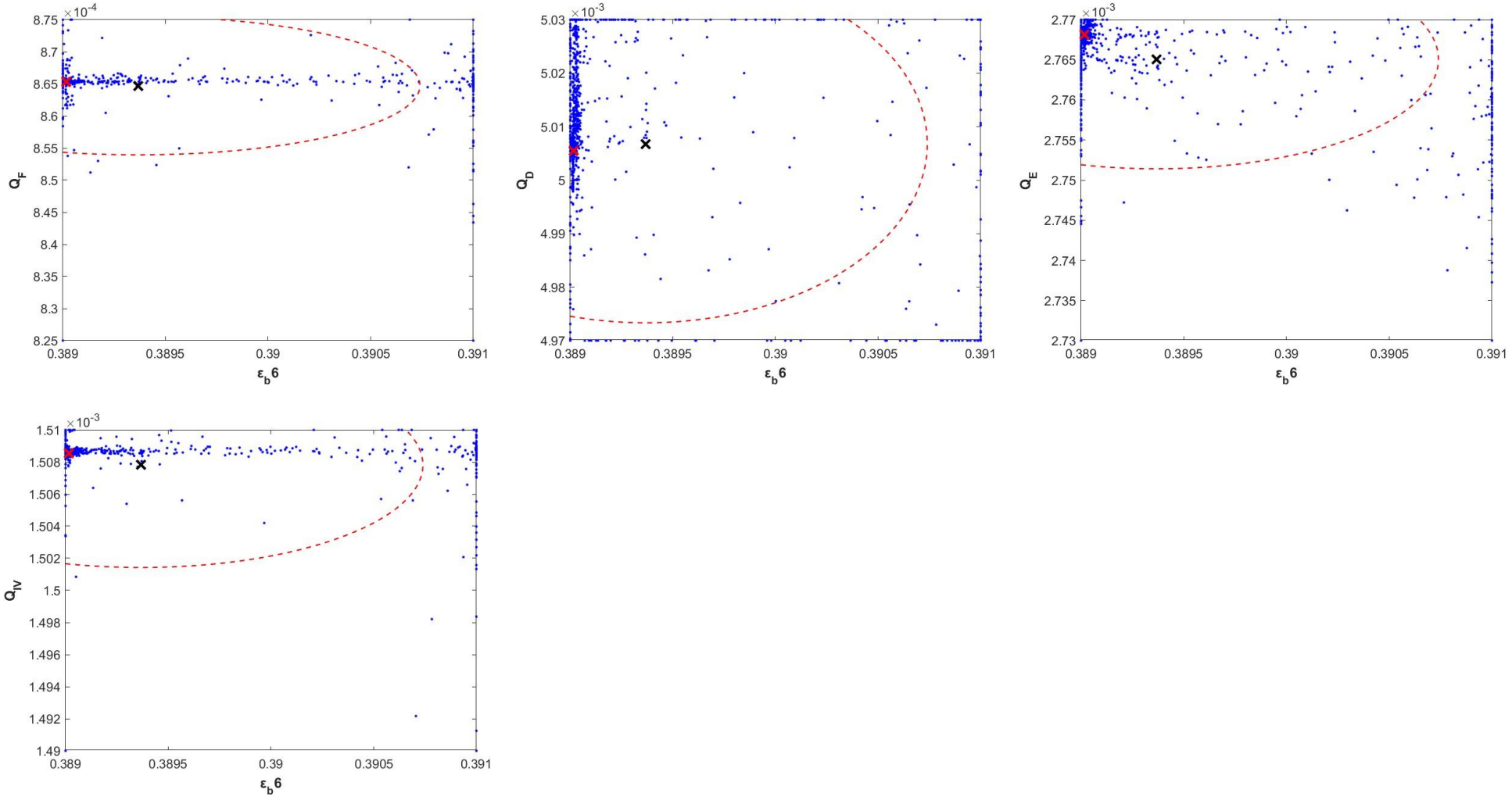

Figure B. 41 - Confidence Regions for each pair of estimable parameters – Pairs containing $\varepsilon_{b,6}$

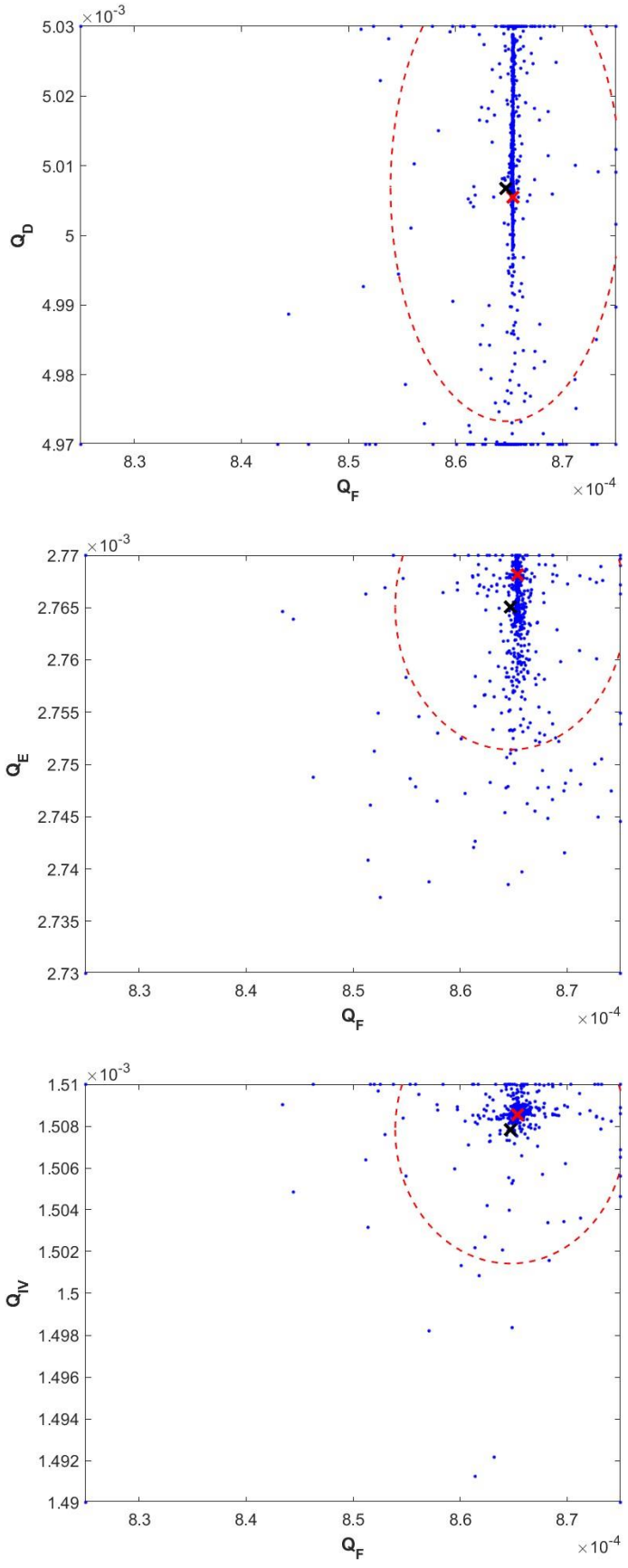

Figure B. 42 - Confidence Regions for each pair of estimable parameters – Pairs containing Q_F

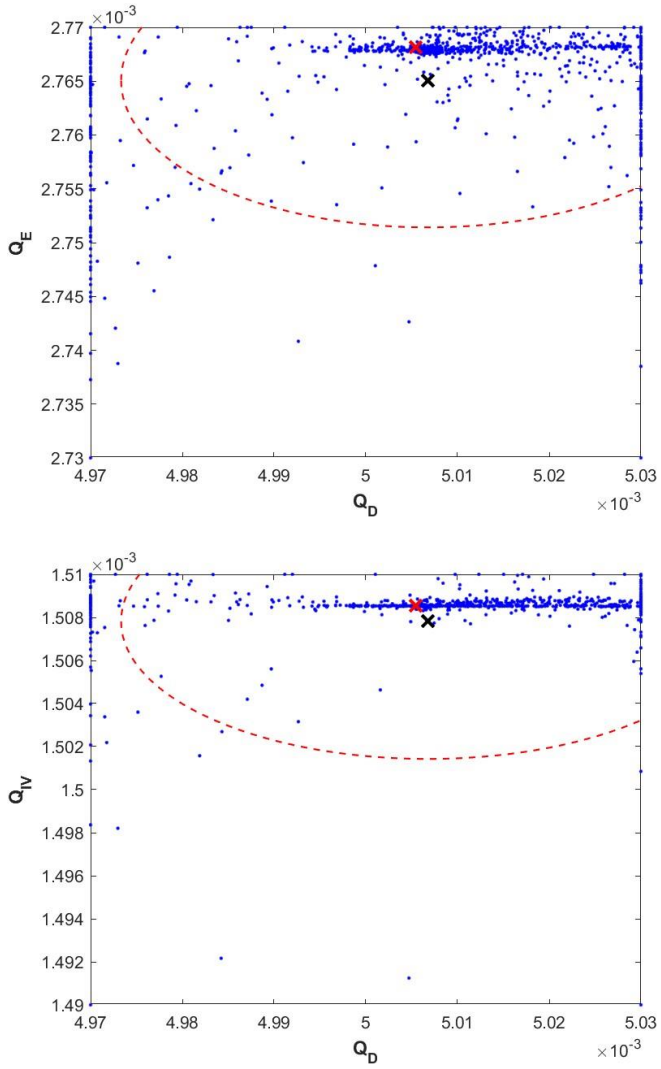

Figure B. 43 - Confidence Regions for each pair of estimable parameters – Pairs containing Q_D

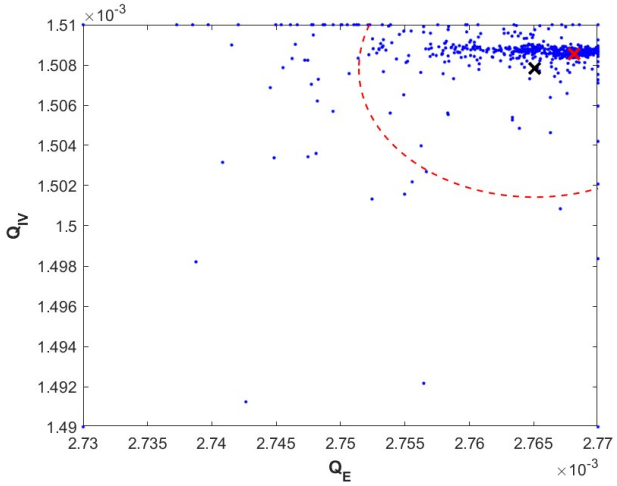

Figure B. 44 - Confidence Regions for each pair of estimable parameters – Pair containing Q_E and Q_{IV}