Master in Chemical Engineering

Influence of Operating Conditions on the precision of the kinetic parameters for different kinetic laws

Master's Thesis

of

David Armando de Oliveira Pinto

Developed for the Course of Dissertation

In

IFP Energies Nouvelles

Supervisor in FEUP: **Prof. José Miguel Loureiro**

Supervisors in IFPen: **Doctor Alberto Silva Servia**

Doctor Jan Verstraete

Chemical Engineering Department

Julho 2016

Acknowledgments

Here I want to express my gratitude to all the people that made this master's thesis, possible.

To Dr. Dominique Humeau, director of the Process Experimentation division, and Dr. Hervé Cauffriez, department head of the Experimentation Intensification department, I thank you for the opportunity to make my internship in IFPEN.

To my supervisors at IFPEN, Doctor Jan Verstraete Doctor Alberto Silva Servia, for all the attention, support, guidance and knowledge given.

To Prof. José Miguel Loureiro, I'm very grateful for his availability, knowledge sharing and all the corrections and important suggestions about this work.

To all my friends here in Lyon, Gonçalo and Paulo, for the support in the hardest moments and for all the time we shared that made this great experience.

To all my family, for their love and support.

To all my friends in Portugal, Eduardo, Daniela and Ana, for all the unconditional support, even hundreds of kilometers away.

To all the people I met during my stay here, for making this ERASMUS experience one of the unforgettable experiences in my life.

Resumo

Na indústria, recursos técnicos e económicos são normalmente limitados. Para determinar qual o modelo cinético para cada reação é então crucial otimizar o número de experiências realizadas.

Neste trabalho, a influência de diferentes condições foram testadas, nomeadamente a função objetivo, modelo cinético e o uso de ruído para a geração de dados pseudo-experimentais.

Utilizando um conjunto de 196 experiências, diferentes métodos para determinar as experiências mais informativas e o melhor método para as determinar.

Deste trabalho, é possível concluir que utilizando menos de oitenta experiências os intervalos de confiança e os valores dos parâmetros não variam significativamente com a adição de mais experiências, permitindo reduzir o seu número para menos de metade do habitual.

Palavras Chave: Modelação cinética, Design Experimental, método D-optimal , método I-optimal

Abstract

In the industrial world, economical and technical resources are often limited. To determine the kinetic model for different reactions it is then crucial to optimize the number of experiments performed.

In this work, the influence of different conditions was tested, such as the objective function, kinetic model and the use of noise to generate pseudo-experimental data.

Using a Data Set of 196 experiments, different methods were tested in order to determine the most informative experiments and the best method to determine them.

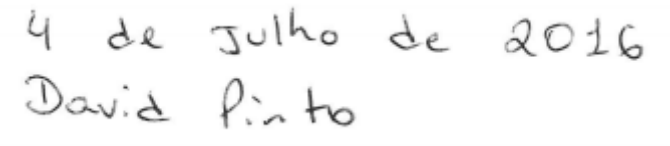
From this work, it is possible to conclude that eighty experiments allow to obtain confidence intervals and parameter values that do not vary significantly after the addition of more experiments. This enables to reduce the number of experiments performed to less than half.

Keywords: Kinetic Modeling, Experimental Design, D-optimal method, I-optimal method

Honor Declaration

I hereby certify and declare on my honor, that the present work is my own and everything coming from external sources has been properly referenced.

Sign and Date



Contents

List of Tables

List of Figures

Nomenclature

y^{exp}	Experimental Conversion	
y^{calc}	Calculated Conversion	
r	Reaction rate	mol/s
k_0	Kinetic constant for the reference temperature	
E_A	Activation Energy	J/mol
R	Ideal gas constant	J/(K.mol)
T	Temperature	K
T_0	Reference Temperature	K
p_{MCH}	Partial pressure of Methyl-Cyclohexane	bar
p_{H_2}	Partial pressure of Hydrogen	bar
n_1	Power law constant for Methyl-Cyclohexane	
n_2	Power law constant for Hydrogen	
K_{MCH}	Adsorption Constant for Methyl-Cyclohexane	
K_{Tol}	Adsorption Constant for Toluene	
y^{p-exp}	Pseudo-Experimental Conversion	

Greek Letters

ε	Noise

Indices

i	Index

List of Abbreviations

MCH	Methyl-Cyclohexane
Tol	Toluene

1 Introduction

1.1 Project Presentation

In the industrial world, economical and technical resources are often limited. To determine the kinetic model for different reactions it is then crucial to optimize the number of experiments performed. Theoretically, the confidence in the obtained kinetic model increases with the number of experiments. The main problems are the limited time and resources to do so.

In this work, the influence of different conditions in the kinetic model is studied.

The objectives of this work are:

- To test two different kinetic models and determine which best represents the experimental data;
- To analyze the evolution of the parameters precision with different objective functions and noise;
- To use an experimental design approach to determine the number of needed experiments to obtain a given parameters precision.

1.2 Company presentation

IFP Energies nouvelles (IFPEN) is a major research and training player in the fields of energy, transport and the environment. From research to industry, technological innovation is central to all its activities.

As part of the public-interest mission with which it has been tasked by the public authorities, IFPEN focuses on:

- providing solutions to take up the challenges facing society in terms of energy and the climate, promoting the transition towards sustainable mobility and the emergence of a more diversified energy mix;
- creating wealth and jobs by supporting French and European economic activity, and the competitiveness of related industrial sectors.

Its programs are structured around 3 strategic priorities:

- **Sustainable mobility**: developing effective, environmentally-friendly solutions for the transport sector;
- **New energies**: producing fuels, chemical intermediates and energy from renewable sources;

- **Responsible oil and gas**: proposing technologies that meet the demand for energy and chemical products while improving energy efficiency and reducing the environmental impact.

An integral part of IFPEN, its graduate engineering school – IFP School – prepares future generations to take up these challenges.

1.3 Work contributions

This work will provide a starting point to determine the most informative experiments to perform in kinetic modeling.

1.4 Thesis Structure

After an introductory Section on the reaction studied, the rate laws, and the optimization and statistical tools, this thesis is composed of 4 Sections that contain my personal work:

- a Section about the influence of the choice of objective function, where a sum of squares of the absolute deviations and a sum of squares of the relative deviation are tested and compared,
- a Section about the choice of the kinetic model, where two different power law models and two different adsorption models are tested and compared,
- a Section about the influence of noise on the experimental data, where fixed relative noises, random relative noises and random Gaussian noises were tested,
- a Section on experimental design, in which different parameter estimations and different experiment selection methods to determine the most informative experiments are compared.

2 State of Art

2.1 Dehydrogenation of Methyl-Cyclohexane

One of the most important reactions in catalytic reforming is the dehydrogenation of six-membered cycloalkanes.

The rate of the endothermic reaction is high enough to impose itself on the other reforming reactions, resulting in a noticeable and significant temperature drop.

The accuracy of temperature profile predictions in adiabatic reactors is highly dependent of the understanding of the kinetics of this dehydrogenation reaction [1].

The dehydrogenation of Methyl-Cyclohexane is going to be used as a reference reaction in this work.

The reaction is shown in figure 1.

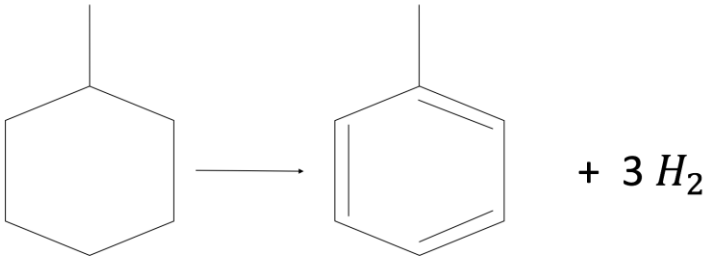


Figure 1 – Methyl-Cyclohexane Dehydrogenation to Toluene

2.2 Experimental Database

The experimental database is composed by experimental points with conditions of:

- Temperature: Between 350 ºC and 400 ºC
- Pressure: Between 6.9 and 17.1bar abs
- H_2 / HC ratio: between 6.9 and 65 mol/mol
- HC flow rate: between 0.13 and 0.30 mol/hr
- Weight Hourly Space Velocity: between 2.23 and 59.2 h^{-1}
- Feed composition: MCH / H_2 mixtures, MCH / H_2 / Tol mixtures

2.3 Types of kinetic laws

2.3.1 Power Law Model

The Power-Law model is a widely used reaction rate model in industry and academia. As shown in equation 2.1, the reaction rate in the power law model is obtained by the product of all the species in the reaction raised to a power.

$$-r = kC_A^a C_B^b \dots \tag{2.1}$$

where r is the reaction rate, k is the kinetic constant, C_A is the concentration of species A, a is the reaction order for specie A, C_B is the concentration of species B and b is the reaction order for specie B.

The individual reaction orders lead to the global reaction order [2].

2.3.2 Adsorption Models

For the reaction of interest, even if the intermediate products are unknown, it is generally accepted that the mechanism consists of a step wise elimination of hydrogen atoms [3]. It was possible to Verstraete [1] to confirm a step wise pathway since he observed, during the performance of experiments, the existence of methyl-cyclohexenes.

The kinetic models used in this work are based on two of the three mechanisms that Verstraete considered during his work. They are represented in figure 2.

Scheme I: Dual site surface reactions with generation of atomic hydrogen

$$MCH + l \rightleftharpoons MCH - l \quad (1)$$

$$MCH - l + l \rightleftharpoons A_1 - l + H - l \quad (2)$$

$$A_1 - l + l \rightleftharpoons MCH^{=} - l + H - l \quad (3)$$

$$MCH^{=} - l + l \rightleftharpoons A_2 - l + H - l \quad (4)$$

$$A_2 - l + l \rightleftharpoons MCH^{2=} - l + H - l \quad (5)$$

$$MCH^{2=} - l + l \rightleftharpoons A_3 - l + H - l \quad (6)$$

$$A_3 - l + l \rightleftharpoons Tol - l + H - l \quad (7)$$

$$Tol - l \rightleftharpoons Tol + l \quad (8)$$

$$2\ H - l \rightleftharpoons H_2 + 2\ l \quad (9)$$

Scheme II: Dual site surface reactions with generation of molecular hydrogen

$$MCH + l \rightleftharpoons MCH - l \quad (10)$$

$$MCH - l + l \rightleftharpoons MCH^{=} - l + H_2 - l \quad (11)$$

$$MCH^{=} - l + l \rightleftharpoons MCH^{2=} - l + H_2 - l \quad (12)$$

$$MCH^{2=} - l + l \rightleftharpoons Tol - l + H_2 - l \quad (13)$$

$$Tol - l \rightleftharpoons Tol + l \quad (14)$$

$$H_2 - l \rightleftharpoons H_2 + l \quad (15)$$

Scheme III: Single site surface reactions according to an Eley-Rideal mechanism

$$MCH + l \rightleftharpoons MCH - l \quad (16)$$

$$MCH - l \rightleftharpoons MCH^{=} - l + H_2 \quad (17)$$

$$MCH^{=} - l \rightleftharpoons MCH^{2=} - l + H_2 \quad (18)$$

$$MCH^{2=} - l \rightleftharpoons Tol - l + H_2 \quad (19)$$

$$Tol - l \rightleftharpoons Tol + l \quad (20)$$

Figure 2 – Reaction mechanisms considered during the work of Verstraete [1]

With these mechanisms twenty different kinetic models were developed [1]. The models considered in this work were the models with controlling step marked in figure 2 with the numbers 12 and 18. All these models are nonlinear and to determine the parameters values that best fit the experimental data it is necessary to use a nonlinear algorithm.

2.4 Optimization for kinetic parameters determination

In this subsection, two different optimization methods will be presented.

2.4.1 Newton's Method

Newton Method uses a quadratic approximation to the function $f(\boldsymbol{x})$. That approximation is shown in equation 2.2.

$$f(\boldsymbol{x}) \approx f(\boldsymbol{x}^k) + \nabla^T f(\boldsymbol{x}^k)\Delta\boldsymbol{x}^k + \frac{1}{2}(\Delta\boldsymbol{x}^k)^T \boldsymbol{H}(\boldsymbol{x}^k)\Delta\boldsymbol{x}^k \tag{2.2}$$

where $f(\boldsymbol{x})$ is the model, $\boldsymbol{x}$ is the independent variable, $\boldsymbol{x}^k$ is a point in the model and $\boldsymbol{H}(\boldsymbol{x}^k)$ is the Hessian matrix of $f(\boldsymbol{x})$.

In kinetics the function $f(\boldsymbol{x})$ is usually dependent of the experimental conditions (Temperature, Concentration…) and the kinetic parameters ($\boldsymbol{\beta}$) being then represented as $f(\boldsymbol{x}, \boldsymbol{\beta})$.

In this particular case, $\boldsymbol{x}^k$ from equation 2.2 is a point of constant parameters.

The minimum of $f(\boldsymbol{x})$ is obtained by differentiating this approximation with respect to each of the components of $\boldsymbol{x}$ and equating the resulting expression to zero. This is represented in equation 2.3.

$$\nabla f(\boldsymbol{x}) = \nabla f(\boldsymbol{x}^k) + \boldsymbol{H}(\boldsymbol{x}^k)\Delta\boldsymbol{x}^k = 0 \tag{2.3}$$

This equation can be represented as shown in equation 2.4.

$$\boldsymbol{x}^{k+1} - \boldsymbol{x}^k = \Delta\boldsymbol{x}^k = -\left[\boldsymbol{H}(\boldsymbol{x}^k)\right]^{-1}\nabla f(\boldsymbol{x}^k) \tag{2.4}$$

Note that if $f(\boldsymbol{x})$ is quadratic only one step is necessary to reach the minimum of $f(\boldsymbol{x})$[4].

2.4.2 Levenberg-Marquardt Method

The Levenberg-Marquardt method is based on Newton's method where the $\boldsymbol{H}(\boldsymbol{x}^k)$ matrix is modified according to equation 2.5.

$$\widetilde{\boldsymbol{H}}(\boldsymbol{x}) = [\boldsymbol{H}(\boldsymbol{x}) + \beta\boldsymbol{I}] \tag{2.5}$$

where $\widetilde{\boldsymbol{H}}(\boldsymbol{x})$ is the modified Hessian matrix, β is an optimization parameter and $\boldsymbol{I}$ is the Identity matrix.

This modified Hessian matrix is replaced in equation 2.3. β is recalculated after each iteration based on figure 3:

$$If\left[\, f(\boldsymbol{x}^{k+1}) < f(\boldsymbol{x}^k) \,\right]$$

$$Then\left[\, \beta^{k+1} = \frac{1}{4}\beta^k \,, and\ k = k+1 \right]$$

$$Else[\ \beta^k = 2\beta^k\ ]$$

Figure 3 - Method to recalculate β [4]

The method used in this work was the Levenberg-Marquardt modification.

2.5 Kinetic parameters statistical analysis

2.5.1 Jacobian matrix

The Jacobian matrix is calculated using equation 2.6.

$$J_{ij} = \frac{\partial f(x_i,\beta)}{\partial \beta_j} \tag{2.6}$$

2.5.2 Hessian Matrix

The Hessian matrix is calculated using equation 2.7.

$$H_{ij} = \frac{\partial^2 f(x_i,\beta)}{\partial x_i \partial x_j} \tag{2.7}$$

2.5.3 Confidence Interval

The confidence interval is calculated by equation 2.8.

$$b_i - s(b_i) \times t_{tab}\left(n-p, 1-\frac{\alpha}{2}\right) \leq \beta_i \leq b_i + s(b_i) \times t_{tab}\left(n-p, 1-\frac{\alpha}{2}\right) \tag{2.8}$$

where β_i is the unknown parameter number i, b_i is the parameter i estimation, $s(b_i)$ is the standard estimation, t_{tab} is the tabulated t-student distribution with $n-p$ being the degrees of freedom and $1-\frac{\alpha}{2}$ the probability to calculate being $100(1-\alpha)\%$ the confidence percentage of the confidence interval [5].

2.5.4 F-Distribution value

The F-distribution value is usually used to compare different models. The F-value for a regression model is calculated by equation 9.

$$F - statistic = \frac{\frac{Regression\ SSQ}{p}}{\frac{Residual\ SSQ}{n-p}} \tag{2.9}$$

2.6 Experimental Design

For this work, the following two different methods were considered.

The D-optimal design is the most used and calculates the best next experiment maximizing the determinant of the information matrix. The determinant is shown in equation 10

$$|\boldsymbol{M}| = \left|\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right| \tag{2.10}$$

where $\boldsymbol{M}$ is the information matrix, $\boldsymbol{X}$ is the experimental conditions matrix and $\boldsymbol{V}^{-1}$ is the inverse of the covariance matrix and $\boldsymbol{X}^{\boldsymbol{T}}$ is the transposed of the $\boldsymbol{X}$ matrix.

The I-optimal design calculates the best next experiment minimizing the average predicted variance. This is calculated by equation 11.

$$Average\ predicted\ variance = \frac{\int_{\gamma}\ f'(\boldsymbol{x})\times\left(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1} f(\boldsymbol{x})\,dx}{\int_{\gamma}\ dx} \tag{2.11}$$

where f is the considered model after parameter linearization, x is experimental conditions and γ is the experimental space.[6]

Both methods work testing for the existent experimental data, and for the number of trials selected, they test which array of experiments maximizes or minimizes the respective objective function.

3 Objective Function

The influence of two different objective functions on the parameters estimation is analyzed in this section.

The different objective functions compared in this chapter are the sum of the squares of the absolute deviations:

$$F_{obj} = \sum_{i=1}^{n} \left(y_i^{exp} - y_i^{calc}\right)^2 \tag{3.1}$$

where y_i^{exp} is the experimental conversion in the experiment i and y_i^{calc} is the conversion calculated by the model for the same experiment, and the sum of the squares of the relative deviations:

$$F_{obj} = \sum_{i=1}^{n} \left(\frac{y_i^{exp} - y_i^{calc}}{y_i^{exp}}\right)^2 \tag{3.2}$$

The optimization was performed considering a single site surface reaction according to Eley-Rideal mechanism, with the generation of the second hydrogen molecule as controlling step mechanism. The mechanism is presented in figure 2 and the controlling step is given by equation number 18.

The kinetic model that describes this is represented in equation 3.3.

$$r = \frac{k_0 \times e^{\frac{E_A}{R} \times \left(\frac{1}{T} - \frac{1}{T_0}\right)} \left(p_{MCH} - \frac{p_{Tol} \times {p_{H_2}}^3}{K_{eq}}\right)}{p_{H_2} \times (1 + K_{MCH} p_{MCH} + K_{Tol} p_{Tol})} \tag{3.3}$$

where r is the reaction rate, T_0 is the reference temperature of 375°C, k_0 is the kinetic constant at the reference temperature, R is the ideal gas constant, T is the temperature, p_{MCH} is the partial pressure of Methyl-Cyclohexane, K_{MCH} is the adsorption constant for Methyl-Cyclohexane in bar^{-1}, p_{H_2} is the partial pressure of Hydrogen, p_{Tol} is the partial pressure of Toluene and K_{Tol} is the adsorption constant for Toluene in bar^{-1}.

The parameter estimation was conducted using the 217 experimental results.

3.1 Sum of the Squares of the Absolute Deviations

The obtained parameters are represented in table 1. The parity plot between measured and predicted values is presented in Figure 4. A relative error of 10 % was assumed for the experimental conversion.

Table 1 -Parameter estimation and confidence interval using the Sum of the Squares of the Absolute Deviations as objective function

Nr.	Parameter. Estimation	Lower Limit	Upper Limit	t-value
k_0	5.519E-01	4.981E-01	6.057E-01	20.52
E_A	2.000E+05	1.916E+05	2.084E+05	47.62
K_{MCH}	4.203E+00	3.599E+00	4.807E+00	13.92
K_{Tol}	1.358E+00	7.671E-01	1.949E+00	4.597

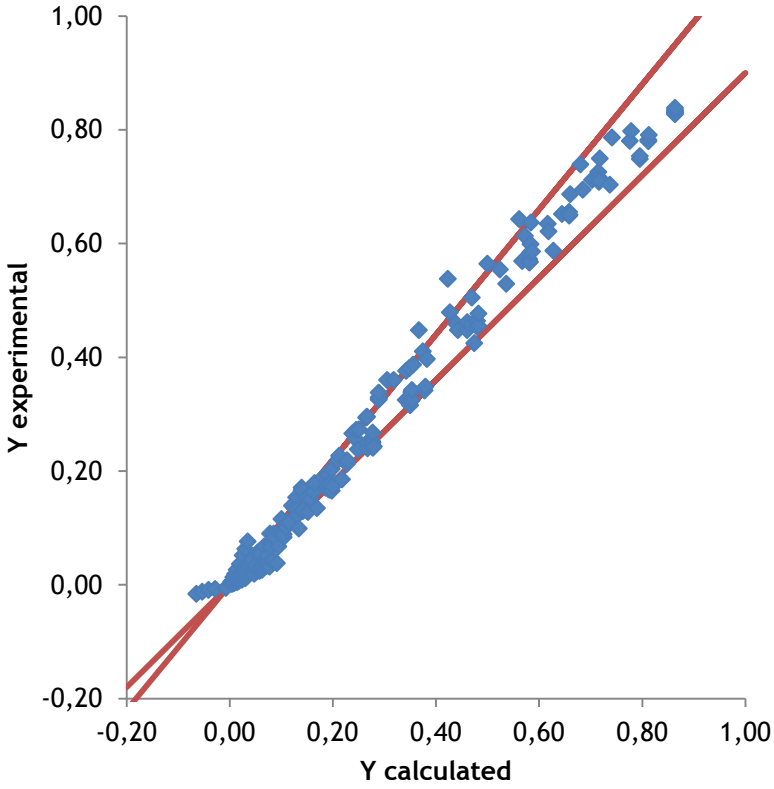


Figure 4 - Parity plot considering an experimental conversion relative error of 0.1 for the Sum of Squares of the Absolute Deviations.

3.2 Sum of the squares of the relative deviations

The obtained parameters are represented in table 2 and the associated parity plot is illustrated in Figure 5. Once again, a relative error of 10 % was considered for the experimental conversion.

Table 2 - Parameter estimation and confidence interval using the Sum of the squares of the relative deviations as objective function

Nr.	Parameter. Estimation	Lower Limit	Upper Limit	t-value
k_0	6.698E-01	4.217E-01	9.179E-01	5.400
E_A	2.599E+05	2.536E+05	2.661E+05	83.71
K_{MCH}	8.756E+00	4.530E+00	1.298E+01	4.144
K_{Tol}	8.150E+00	4.205E+00	1.210E+01	4.132

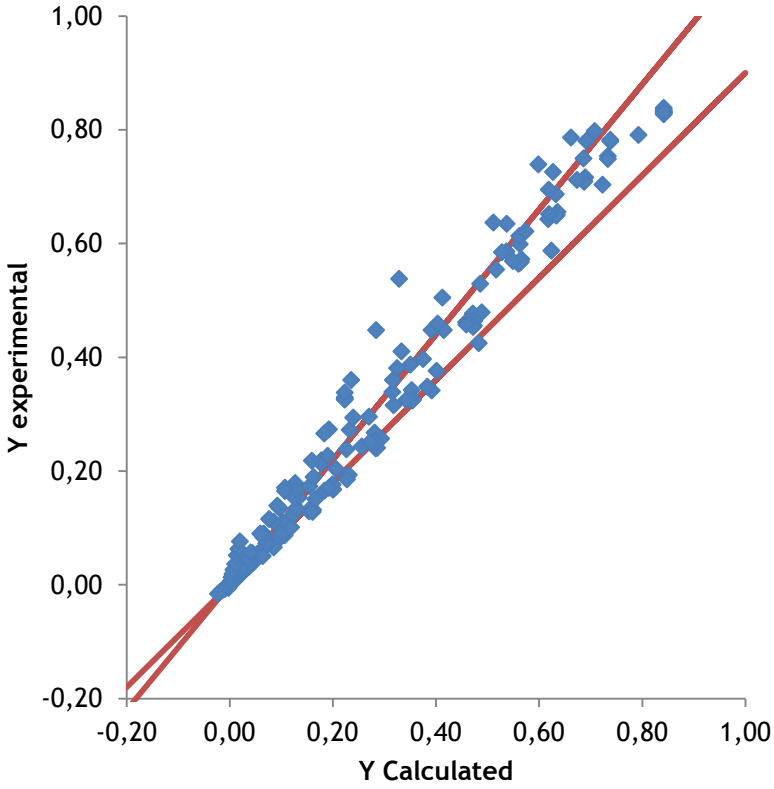


Figure 5 - Parity plot considering an experimental conversion relative error of 0.1 for the Sum of Squares of the Relative Deviations

3.3 Conclusions

It is possible to observe that the sum of the squares of the absolute deviations allows the estimation of the higher conversions to be enhanced (see Figure 4), while the sum of the squares of the relative deviations shows better results at the lower conversions (see Figure 5). There are only a few experimental points obtained at relatively low conversion, and as a result, the conclusion is that for this particular case, the sum of the squares of the absolute deviations is more suitable. Moreover, under industrial operating conditions, very high conversions are observed. Hence, the accurate prediction of high conversions is considered to be more important than the accurate prediction of low conversions.

4 Kinetic Model

Different Kinetic Models were tested in this Section with the objective to determine the best fitting model to this reaction.

The optimization was conducted using the Sum of the Squares of the Absolute Deviations (equation 3.1) as the objective function.

4.1 Power Law

Two different power law models were first of all used for data interpretation. The first one considers the influence of MCH, H2 and TOL, considering the equilibrium reaction, and the second one the impact of TOL is considered null.

The second tested power law model is shown in Appendix A.

4.1.1 5 Parameters Power Law

The first power law considered was:

$$r = k_0 \times e^{\frac{E_A}{R}\times\left(\frac{1}{T}-\frac{1}{T_0}\right)} \times {p_{MCH}}^{n_1} \times {p_{H_2}}^{n_2} \times {p_{Tol}}^{n_3} \quad (4.1)$$

where n_1 is the order for Methyl-Cyclohexane, n_2 is the order for Hydrogen and n_3 is the order for Toluene.

The obtained parameters and the corresponding confidence intervals and t-values are presented in Table 3.

Table 3 - Parameter estimation and confidence interval for a Power Law considering the equilibrium reaction

Nr.	Parameter. Estimation	Lower Limit	Upper Limit	t-value
k_0	9.111E-02	6.479E-02	1.174E-01	6.923
E_A	2.233E+05	2.106E+05	2.359E+05	35.31
n_1	5.530E-01	5.035E-01	6.026E-01	22.32
n_2	-1.207E+00	-1.321E+00	-1.093E+00	-21.21
n_3	-2.228E-01	-2.569E-01	-1.886E-01	-13.04

The parity plot for this model is shown in figure 6

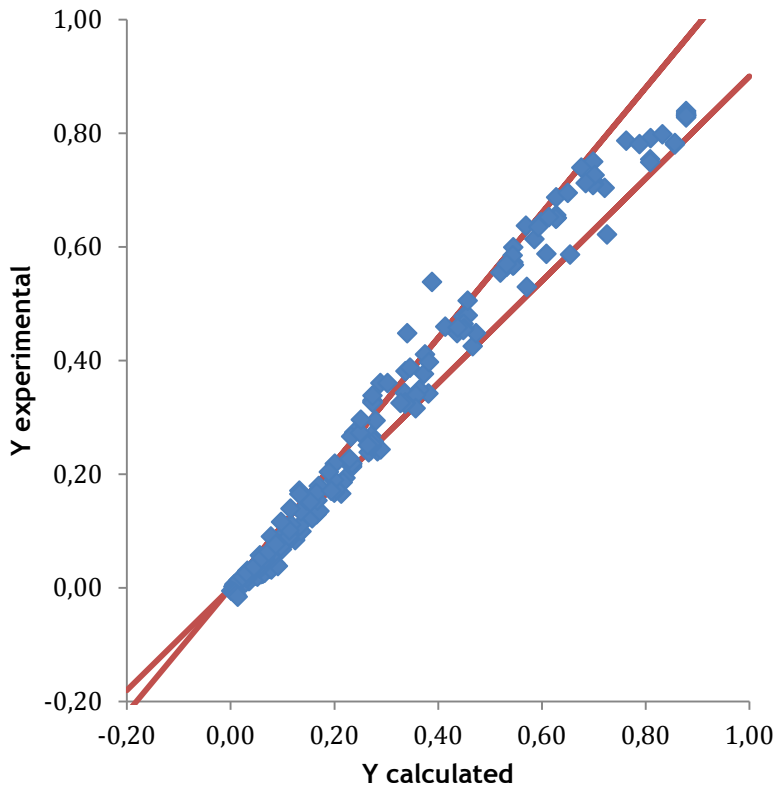


Figure 6 - Parity plot considering an experimental conversion relative error of 0.1 for a Power Law with 5 parameters

The F-value obtained for this model was 5215.133.

4.2 Adsorption Models

Different Langmuir-Hinshelwood kinetic models were tested for this reaction. The considered models were taken from the thesis of Verstraete [1]. The different models are described in the next sub-chapters.

4.2.1 Dual site surface reactions with generation of molecular hydrogen, with the generation of the second hydrogen molecule as controlling step

The first model that was studied is the model that considers dual site surface reactions with generation of molecular hydrogen, with the generation of the second hydrogen molecule as controlling step. The reaction rate equation is represented in equation 4.2.

$$r = \frac{k_0 \times e^{\frac{E_A}{R} \times \left(\frac{1}{T} - \frac{1}{T_0}\right)} \times \left(p_{MCH} - \frac{p_{Tol} \times {p_{H_2}}^3}{K_{eq}}\right)}{p_{H_2} \times (1 + K_{MCH} p_{MCH} + K_{Tol} p_{Tol})^2} \tag{4.2}$$

The obtained parameters are shown in table 4. The parity plot for this model is shown in figure 7.

Table 4 - Parameter estimation and confidence interval for dual site surface reactions with generation of molecular hydrogen, with the generation of the second hydrogen molecule as controlling step

Nr.	Parameter. Estimation	Lower Limit	Upper Limit	t-value
k_0	3.992E-01	3.716E-01	4.269E-01	28.88
E_A	2.025E+05	1.934E+05	2.115E+05	44.77
K_{MCH}	9.568E-01	8.702E-01	1.043E+00	22.11
K_{Tol}	2.703E-01	1.262E-01	4.144E-01	3.752

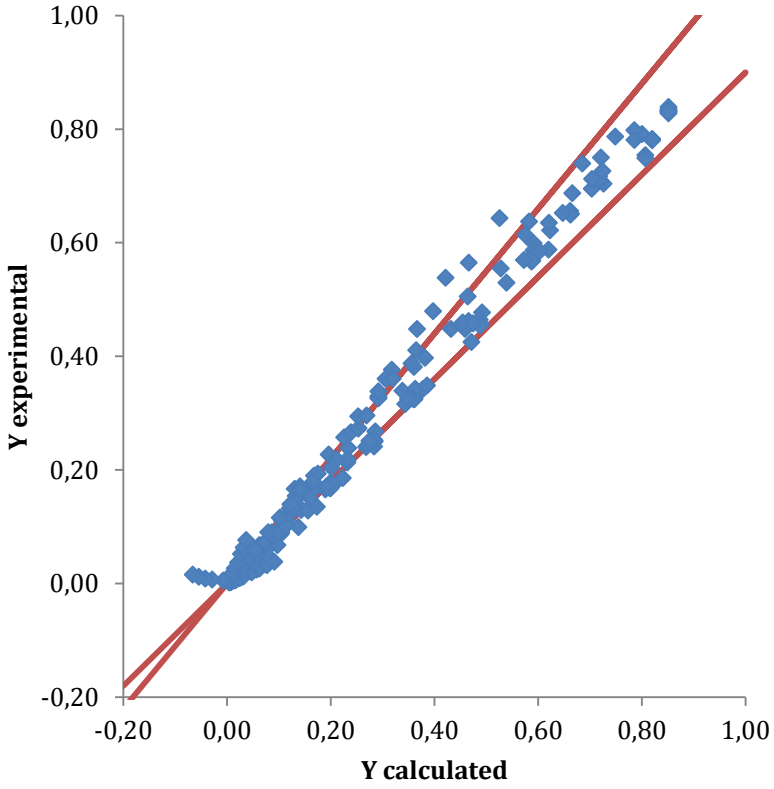


Figure 7 - Parity plot considering an experimental conversion relative error of 0.1 for a Dual site surface reactions with generation of molecular hydrogen, with the generation of the second hydrogen molecule as controlling step model

The F-value obtained for this model was 8983.873

4.2.2 Single site surface reactions according to the Eley-Rideal mechanism, with the generation of the second hydrogen molecule as controlling step

The second model that was studied is the model that considers single site surface reactions according to Eley-Rideal mechanism, with the generation of the second hydrogen molecule as controlling step. The reaction rate equation is represented in equation 4.3.

$$r = \frac{k_0 \times e^{\frac{E_A}{R} \times \left(\frac{1}{T} - \frac{1}{T_0}\right)} \times \left(p_{MCH} - \frac{p_{Tol} \times {p_{H_2}}^3}{K_{eq}}\right)}{p_{H_2} \times (1 + K_{MCH} p_{MCH} + K_{Tol} p_{Tol})} \quad (4.3)$$

The obtained parameters are shown in table 1, and the parity plot for this model are shown in figure 4 both in chapter 3.

The F-value obtained for this model was 10473.167.

4.3 Conclusions

Comparing the F-value for every considered model, it is possible to conclude that the best fit is the model with single site surface reaction with the generation of the second hydrogen molecule as the controlling step.

This statistical analysis allows the identification of the most suitable model through the evaluation of the F-value. Very often, a power law kinetic model is chosen to represent experimental data, such as for example the one containing the contribution of toluene. The fit between experimental and model data is quite acceptable (see Figure 6), however this type of model does not provide insight on the real mechanism. Thus, most of the times, the adsorption-based kinetic models are more suitable than the power law ones, unless the experimental database is not sufficient to ensure a good precision for all parameters. In this case, a power law would be preferred, but special care should be taken when using this type of approach when extrapolating the results.

5 Noise Influence in Parameter Estimation

The confidence interval evolution for all the kinetic parameters of the most suitable model (see Section 4) was studied as a function of several types of noise artificially introduced in numerically generated data.

The methodology used in this section is described in the following subchapter. Using the best fitting model, the conversions for the data points were predicted and, after that, different types and values of errors were applied to the calculated conversions in order to generate pseudo-experimental data. These data were then optimized with the same method used in the previous chapters.

5.1 Fixed Relative Noise

The first type of noise studied was a fixed relative noise applied to the different data points. This is described by equation 5.1.

$$y^{p-exp} = (1 \pm \varepsilon) \times y^{calc} \tag{5.1}$$

where y^{p-exp} is the pseudo-experimental conversion for the different noise values, ε is the noise value and y^{calc} is the conversion calculated by the model.

In this case, a fixed ratio of positive and negative signals was used in the applied noise.

The variation of the parameters and the confidence intervals are shown in figure 8.

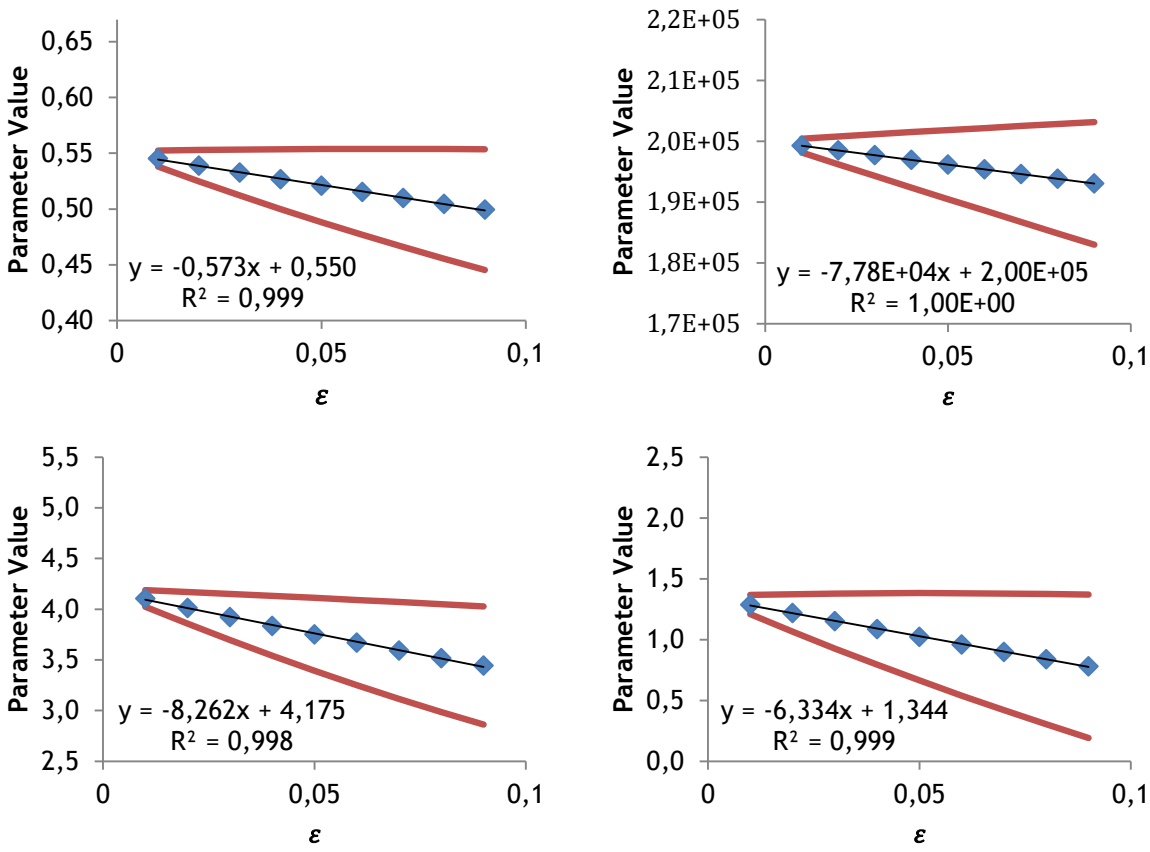


Figure 8 - Parameters estimation for the different fixed noise values where a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

The extrapolated value for a noise of 0 (zero) is near the value of the true parameter, which was expected since the pseudo-experimental data is generated by the model.

A 52.53% of positive signals was considered. The expected results were that the parameter estimation would be kept more or less constant. That did not happen. The main motive that may influence this is that the negative signals are mostly concentrated in the higher conversions.

The almost linear variation in the parameter estimation and in the confidence intervals was expected since the fixed noise signals imply that the only factor changing is the noise value.

The variation of the t-value is shown in figure 9.

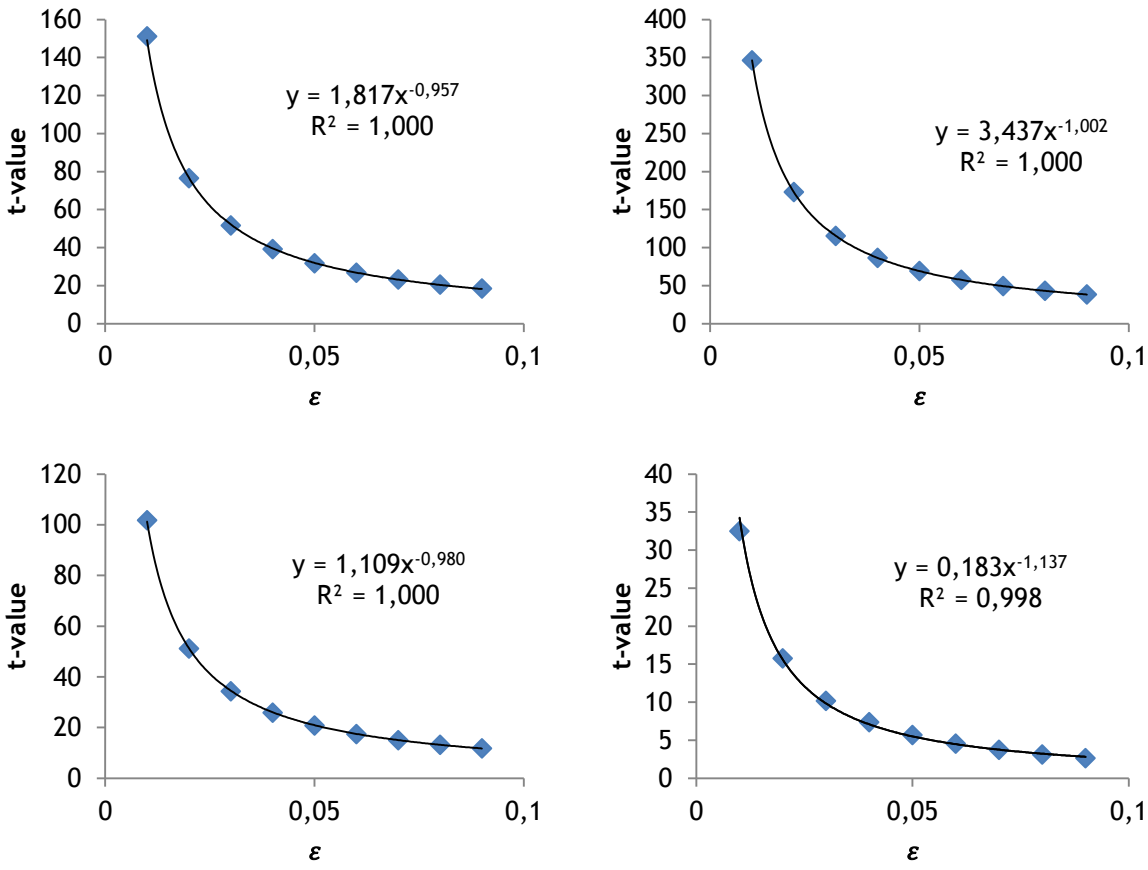


Figure 9 - t-value for the different noise values in which a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

It is observed that the t-value is well represented by a Power regression, having a correlation factor of approximately 1.

5.2 Random Relative Noise

The second type of noise studied was a random relative noise applied to the different data points. This is described by equation 5.2.

$$y^{p-exp} = y^{calc} \times [1 + \varepsilon \times (2 \times Rand() - 1)] \tag{5.2}$$

where Rand() is the excel function that generates a random number between 0 and 1.

In this subsection the influence of having or not a specific number of positive and negative signals is assessed.

In figure 40 in, Appendix B, the variation of the parameters and the confidence intervals for fixed ratios of signals are shown, and in figure 42, in Appendix B, for random ratios of signals.

The same ratio in positive and negative signs was used in this case (and in the same conversion points as in the last one).

The main difference between the fixed noise and this case is that in this subsection the noises are always random, in all data. That makes the parameter calculated for the conversions with noise to stabilize around the true parameter value.

Considering that, in this case, not only the noise value is random, but the signal ratio too, this makes that the parameter estimation is more unstable in this case.

The variation of the t-value for fixed signals ratio is shown in figure 41, in Appendix B, and figure 43, in Appendix B, for random signals ratio.

As it is possible to observe, the t-value variation with the noise is still well represented by a power regression but has some little deviations comparing to the fixed noise.

In figure 10 are represented both t-values for the fixed and random signals ratio.

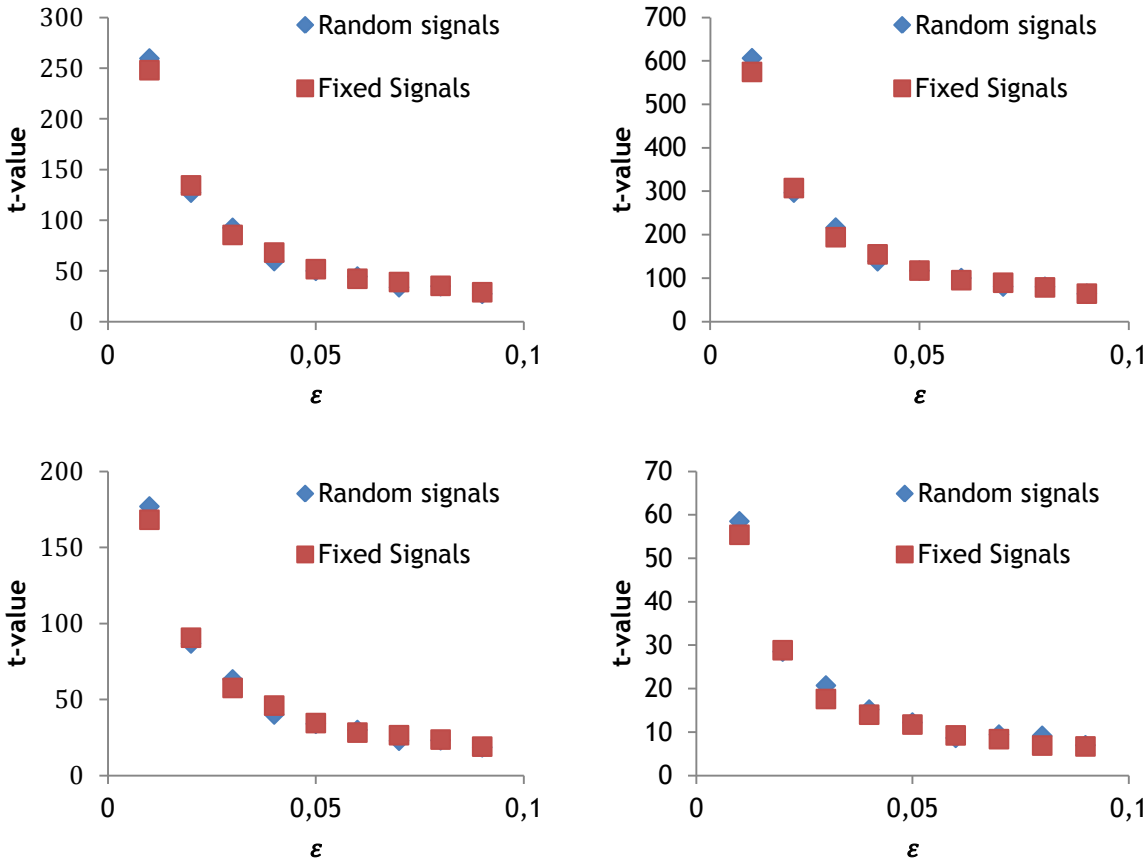


Figure 10 - t-value comparison for the different random noise values with a fixed and random signal ratio where a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

From the analysis of the parameters plots it is possible to conclude that the ratio of positive and negative signals influences the value of the parameter, as it is expected, but from the analysis of figure 11 it is obvious that the t-value is not affected that much by the ratio of signals.

5.3 Gaussian Error Distribution

The third type of noise studied was a Gaussian relative noise applied to the different data points. This is shown in equation 5.3.

$$y^{p-exp} = y^{calc} \times [1 + NormInv(Rand(),0,\varepsilon)] \quad (5.3)$$

where NormInv() is the excel function that returns the inverse of a Gaussian distribution.

In this case it was used a fixed ratio of positive and negative signals in the applied noise.

In figure 44, in Appendix B, the variation of the parameters and the confidence intervals are shown.

As seen in section 5.2, the variation of the parameters and of the confidence intervals depends of the calculated different pseudo-experimental conversions. It can be seen, near a noise value of 0.07, that the non-fixed noise proportion and signals in the generation of the pseudo-experimental data may cause an abrupt change in the value of the parameter. It is possible that this is originated because of the random nature of the noise, since the noise can take any value.

The variation of the t-value is shown in figure 45, in Appendix B.

It is possible to observe that the t-value variation with the noise is still well represented by a power regression but has some little deviations comparing to the fixed noise.

5.4 Conclusions

From t-values it is possible to conclude that, from the applied noises, the one that has the lowest influence on the results is the random relative noise. The implementation of this noise in MsExcel is relatively easy with MsExcel Random function.

Both fixed relative noises and the random relative Gaussian noise give approximately the same t-value, therefore the confidence intervals are approximately the same.

Looking at the parameters, the random relative noise gives parameters always near the true parameter values.

6 Experimental Design

In industry, the technical and economic resources are often limited, as well as the experimental databases. Thus, it is essential to know what are the most informative experiments to be selected in order to develop robust kinetic models. The experimental design approach was used in this work in order to generate the experimental dataset allowing a better precision on the kinetic parameters estimation. Two different approaches were used to sequentially select the most informative experimental points.

The first method, described in section 6.1, was performed using different initial parameters estimations. The Jacobian matrix was then constructed and used to calculate each new experiment.

The second method, described in section 6.2 differed in the calculation of the next experiment, as it performed the parameters estimation at each new experiment selection to recalculate the Jacobian matrix, which is then used to determine a next experiment.

The used model was the Single site surface reactions according to Eley-Rideal mechanism, with the generation of the second hydrogen molecule as controlling step.

In addition to the previous used programs, two different ones were used in this section. These two programs were implemented in R language, using the program RStudio as graphic interface and using optFederov included in AlgDesign library [7].

For an easier implementation, two different routines in MsExcel were developed to sequentially run the different programs and convert file formats between different programs.

6.1 Experimental Design using the Jacobian Matrix Calculated using the Initial Parameter Estimation

Different initial parameters estimations were considered.

6.1.1 Model parameters using D-optimal method

The first considered parameters were the parameters obtained in section 4.2.2. These parameters are shown in table 5.

Table 5 – Initial Estimation Value

Parameter	Estimated value
k_0	5.519E-01
E_A	2.000E+05
K_{MCH}	4.203E+00
K_{Tol}	1.358E+00

Figures 11 to 14 show the parameter estimation, the confidence intervals and the t-value.

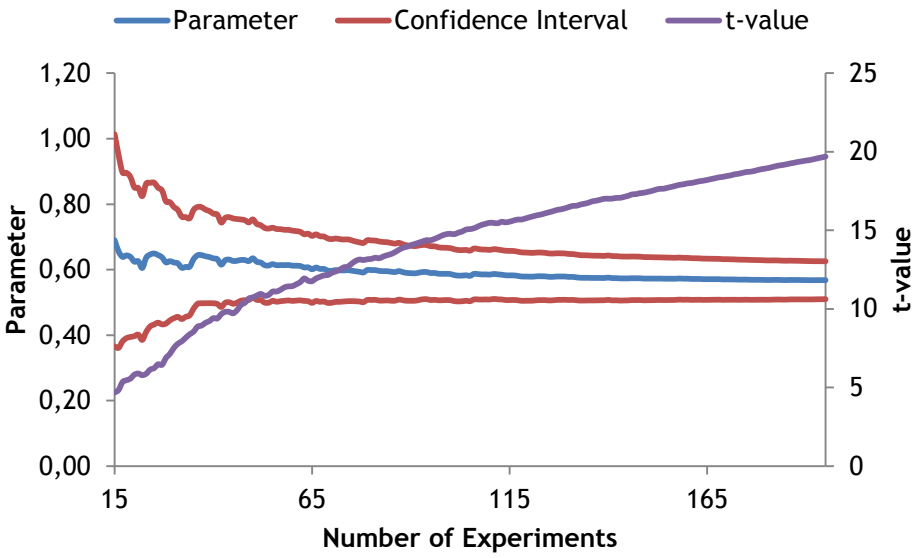


Figure 11 - k_0 Parameter estimation, confidence interval and t-value

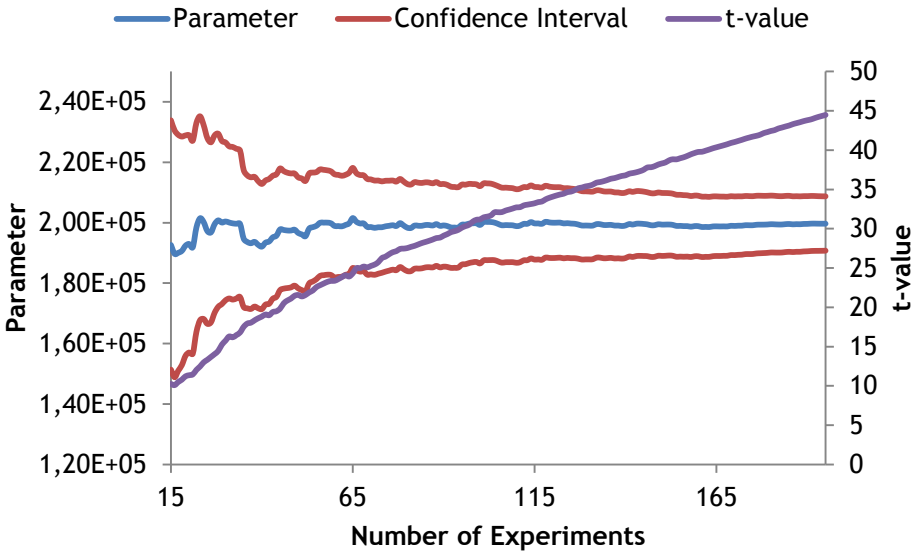


Figure 12 - E_A Parameter estimation, confidence interval and t-value

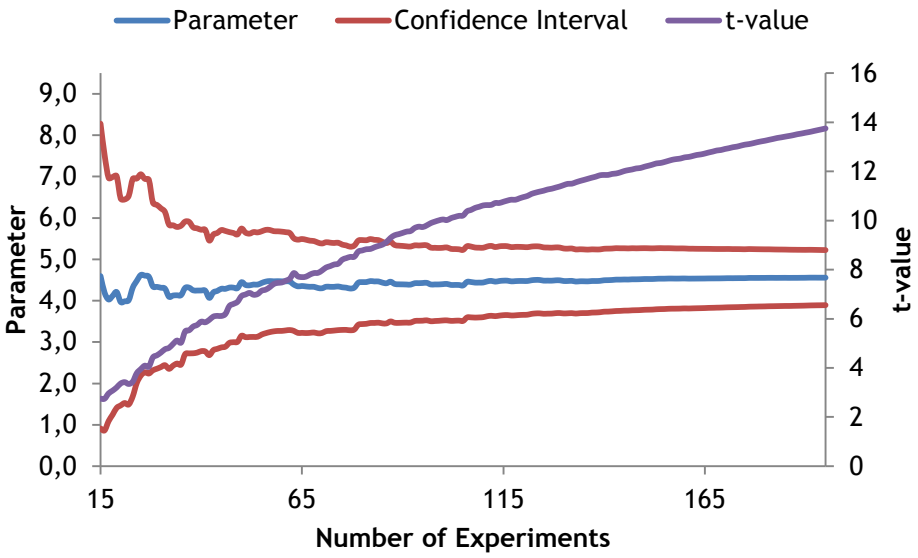


Figure 13 - K_{MCH} Parameter estimation, confidence interval and t-value

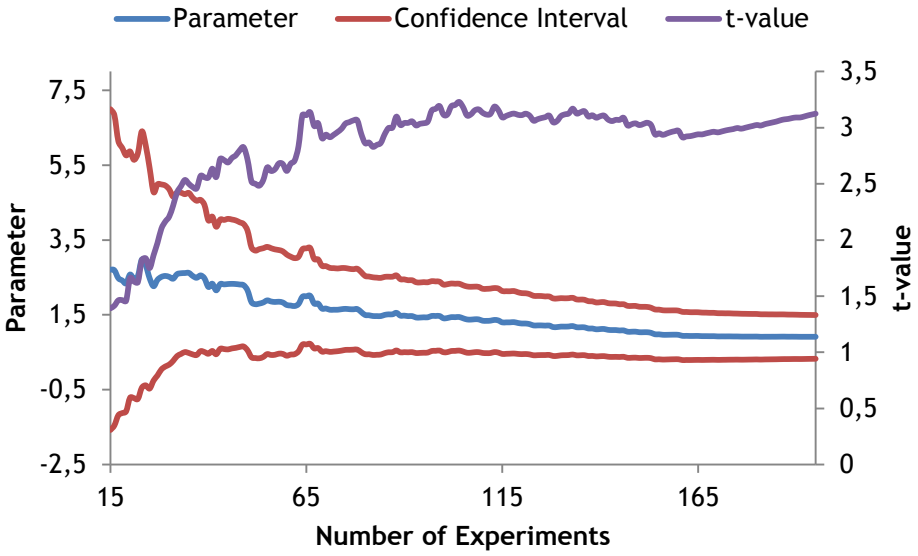


Figure 14 - K_{Tol} Parameter estimation, confidence interval and t-value

As expected, with a small number of experiments the confidence intervals variation is larger when a newer experience is added (in comparison with a large number of experiments). After a number of 70 to 100 experiments the parameters variation and the confidence intervals are so small that each experience added in the top of the previous ones, does not contribute with a lot of information.

For K_{Tol} confidence intervals and the parameter value are more unstable than for the other cases since the experimental dataset has a low number of data points with TOL pressure different of 0 (zero).

For an easier analysis of the confidence interval, the plot of the difference between the upper and lower bound was drawn. That plot is represented in figure 15.

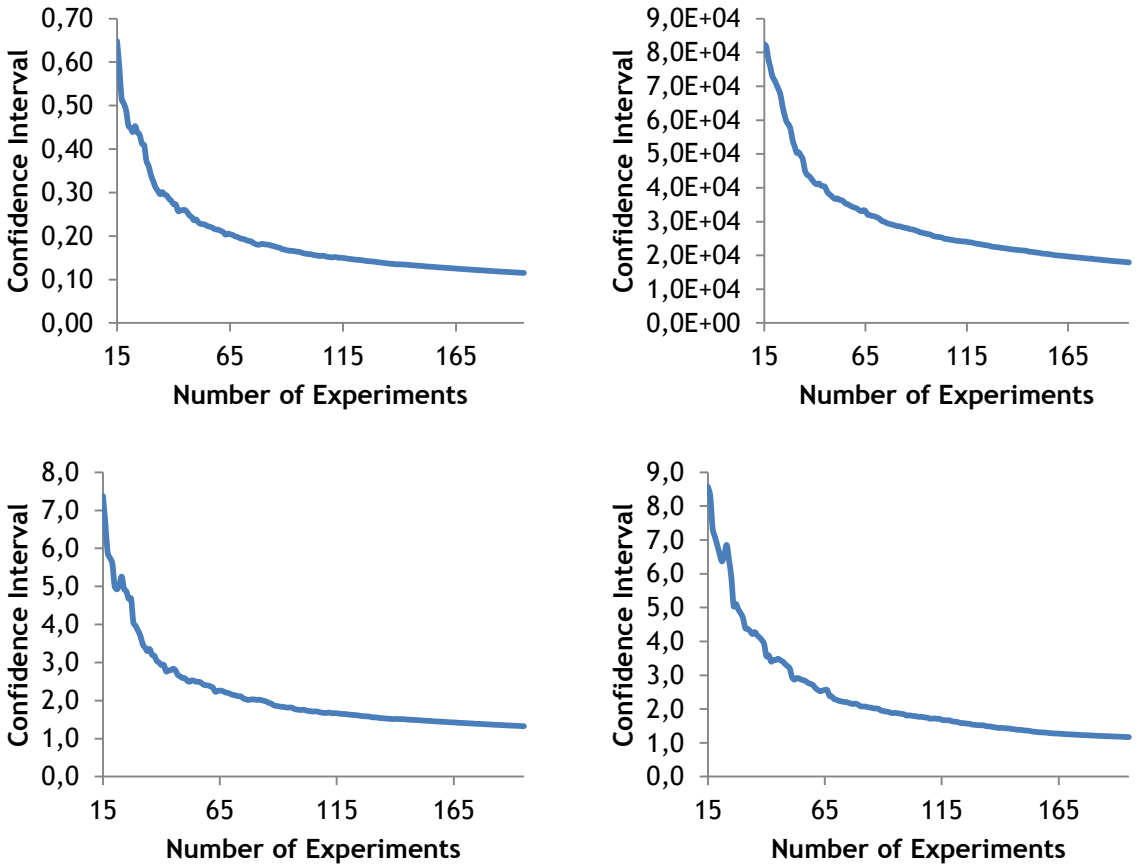


Figure 15 - Confidence Interval Bounds difference for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

In this figure, it is easier to see the variation of the confidence interval as we add experiments and corroborate what was explained before.

In this particular figure we can see that one of the experiments (experience number 23) produces an increase of the confidence interval.

The variation of the parameters calculated after each experiment is represented in figure 16.

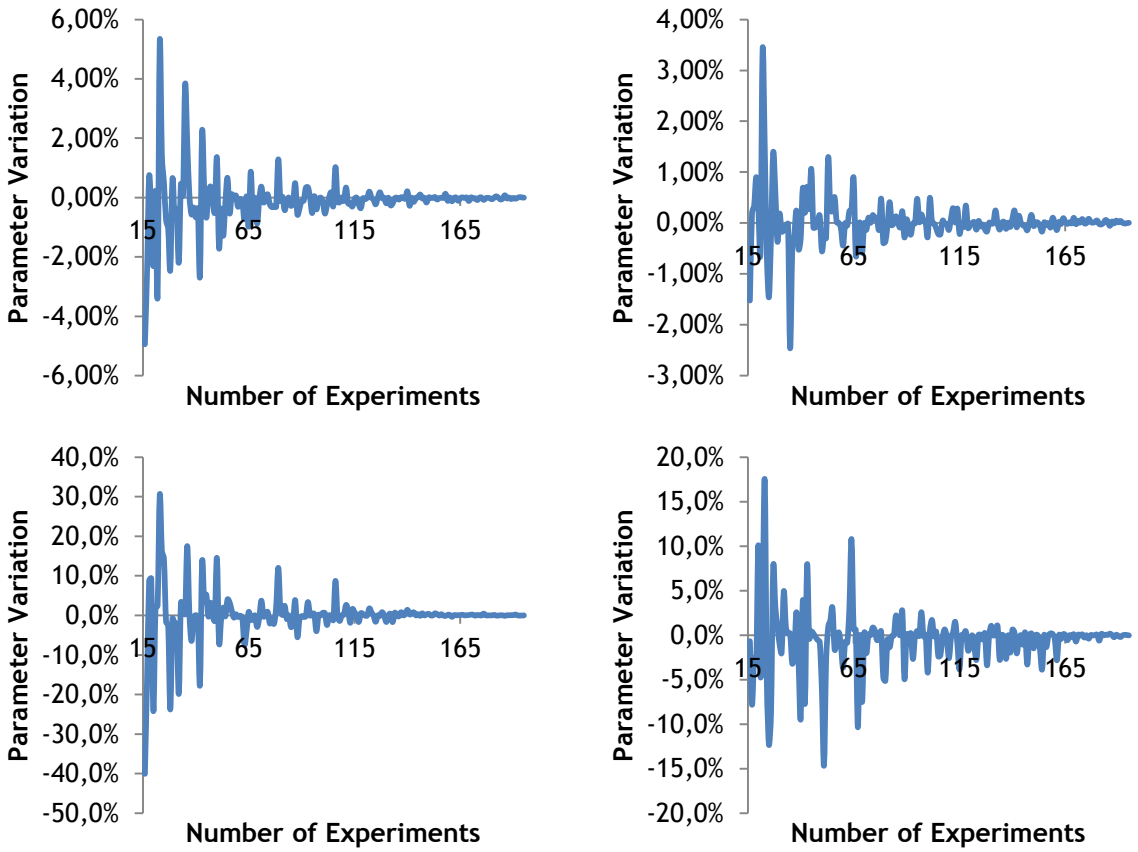


Figure 16 - Parameter variation with each experiment performed for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

In this case, K_{MCH} starts variating less than 10% after experiment number 79.

This corroborates that all the experiments added after experiment number 79 do not add significant information to the parameter value.

The variation of the Confidence Intervals bounds after each experiment is presented in figure 17.

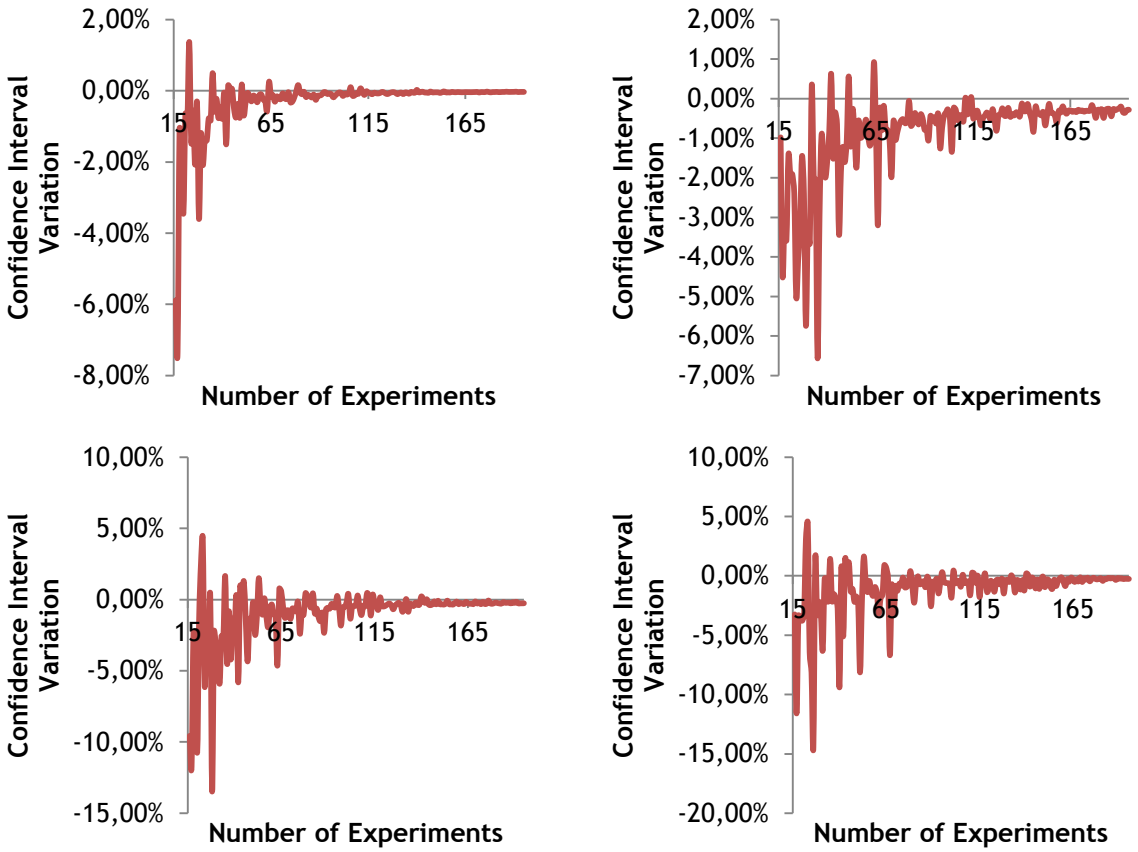


Figure 17 - Confidence Interval bounds variation with each experiment performed for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

Despite the variation of the parameter value is lower than 10% after experiment 79, the confidence interval has a variation lower than this after experiment 28. That means that after this experiment the value of the confidence interval bounds has the biggest variation because of the parameter estimation and not because of the influence of the experiments in this confidence interval.

6.1.2 Plausible estimations using D-Optimal

The second considered parameters are shown in table 6.

Table 6 - Initial Estimation Value

Parameter	Estimated value
k_0	1.000E+00
E_A	1.000E+05
K_{MCH}	1.000E+00
K_{Tol}	1.000E+00

Figures 18 to 21 show the parameter estimation, the confidence intervals and the t-value.

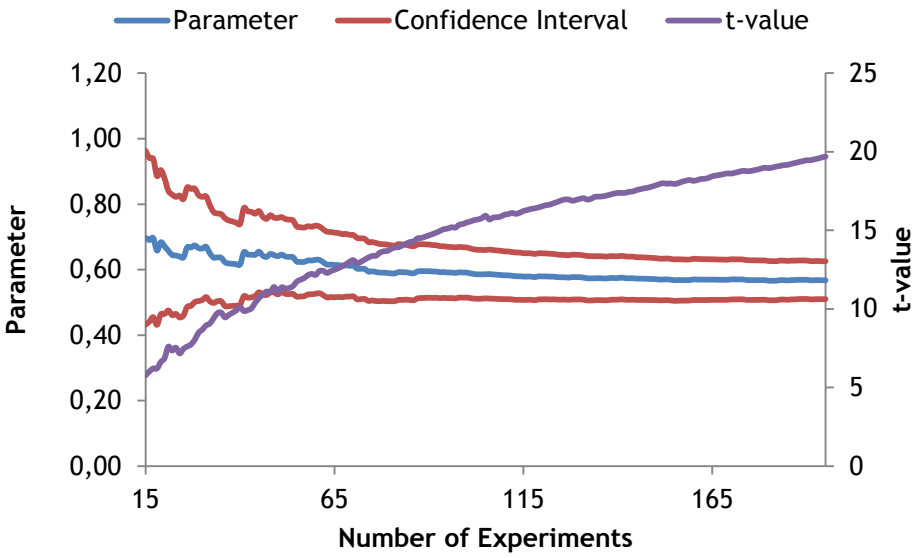


Figure 18 - k_0 Parameter estimation, confidence interval and t-value

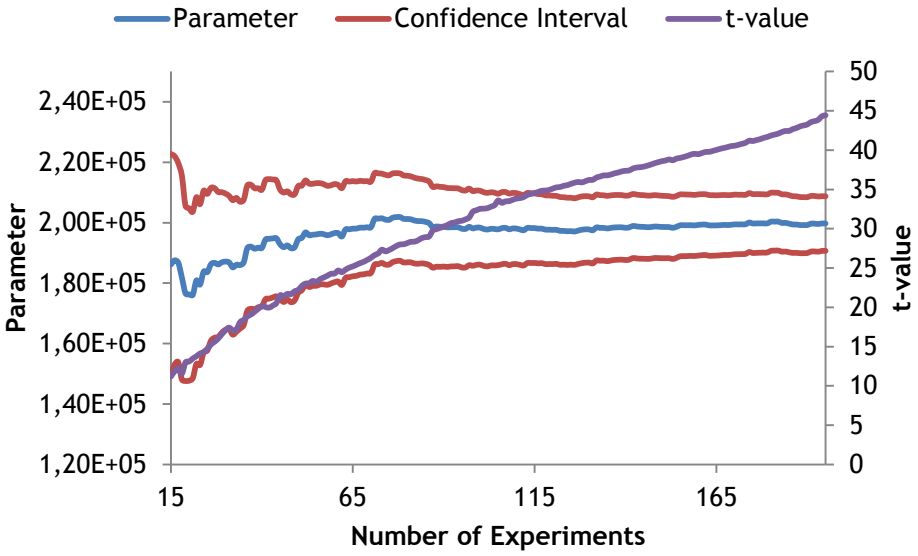


Figure 19 - E_A Parameter estimation, confidence interval and t-value

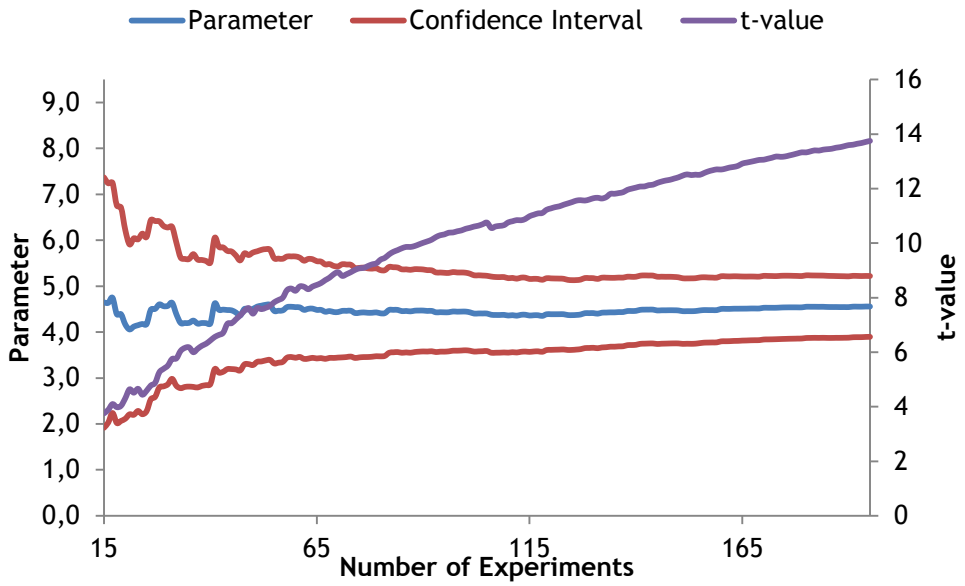


Figure 20 - K_{MCH} Parameter estimation, confidence interval and t-value

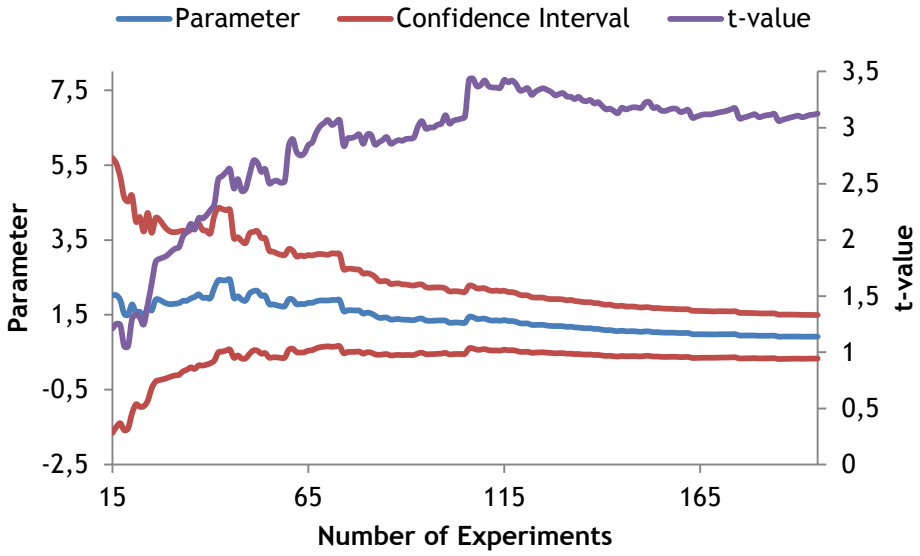


Figure 21 - K_{Tol} Parameter estimation, confidence interval and t-value

Comparing the figures, it is possible to conclude that, using this Jacobian matrix, the parameter estimation after each experiment is much more unstable than in the previous case.

For an easier analysis of the confidence interval, the plot of the difference between the upper and lower bound was drawn. That plot is represented in figure 22.

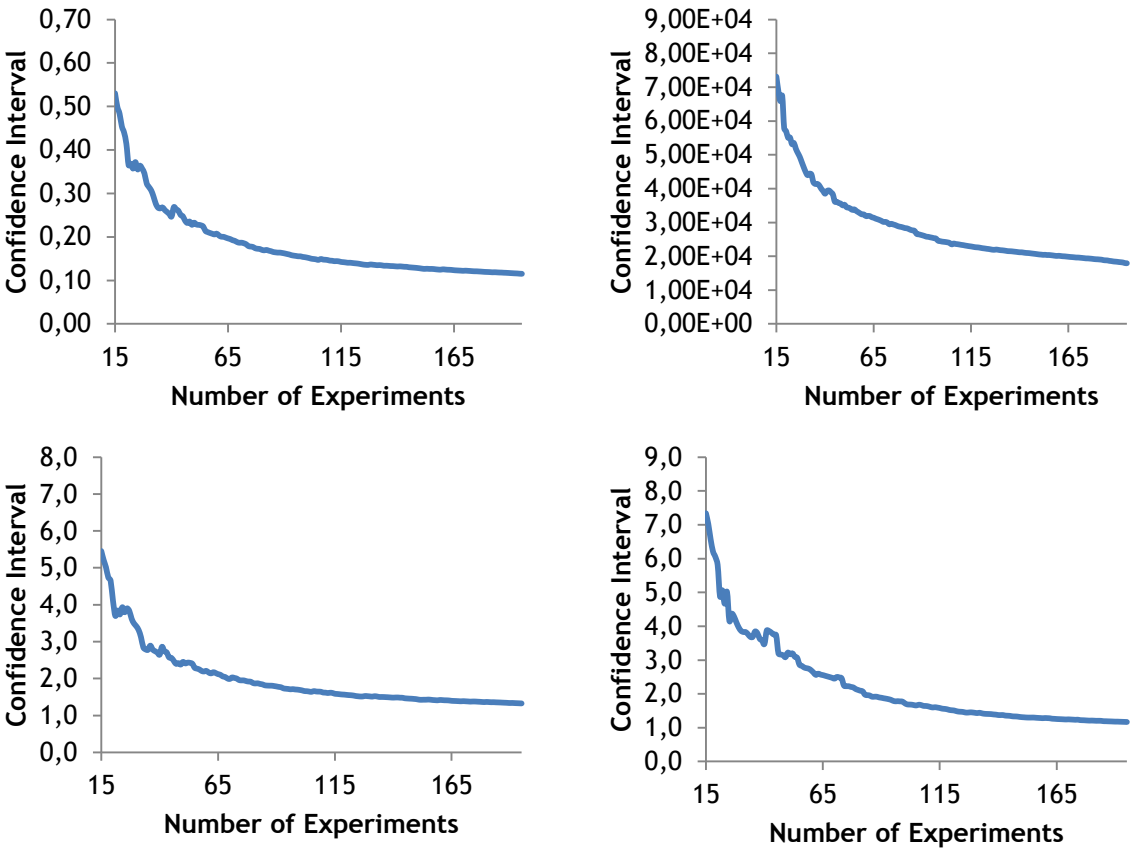


Figure 22 - Confidence Interval Bounds difference for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

The same instability is shown in the Confidence Intervals. The variation of the parameters calculated after each experiment is represented in figure 23.

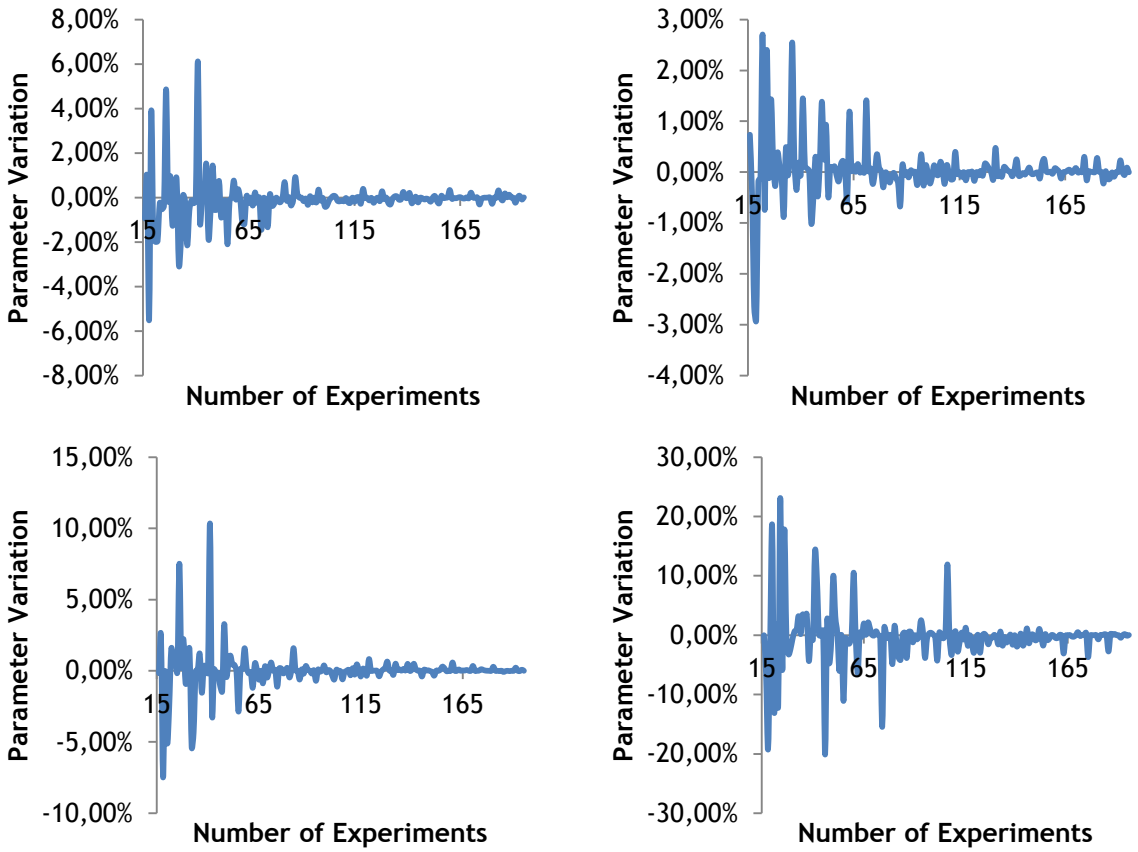


Figure 23 - Parameter variation with each experiment performed for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

In this case, K_{Tol} varies less than 10% after experiment 106. That shows that this initial parameters estimation is less robust than the considered in the previous one.

The variation of the Confidence Intervals bounds after each experiment is represented in figure 24.

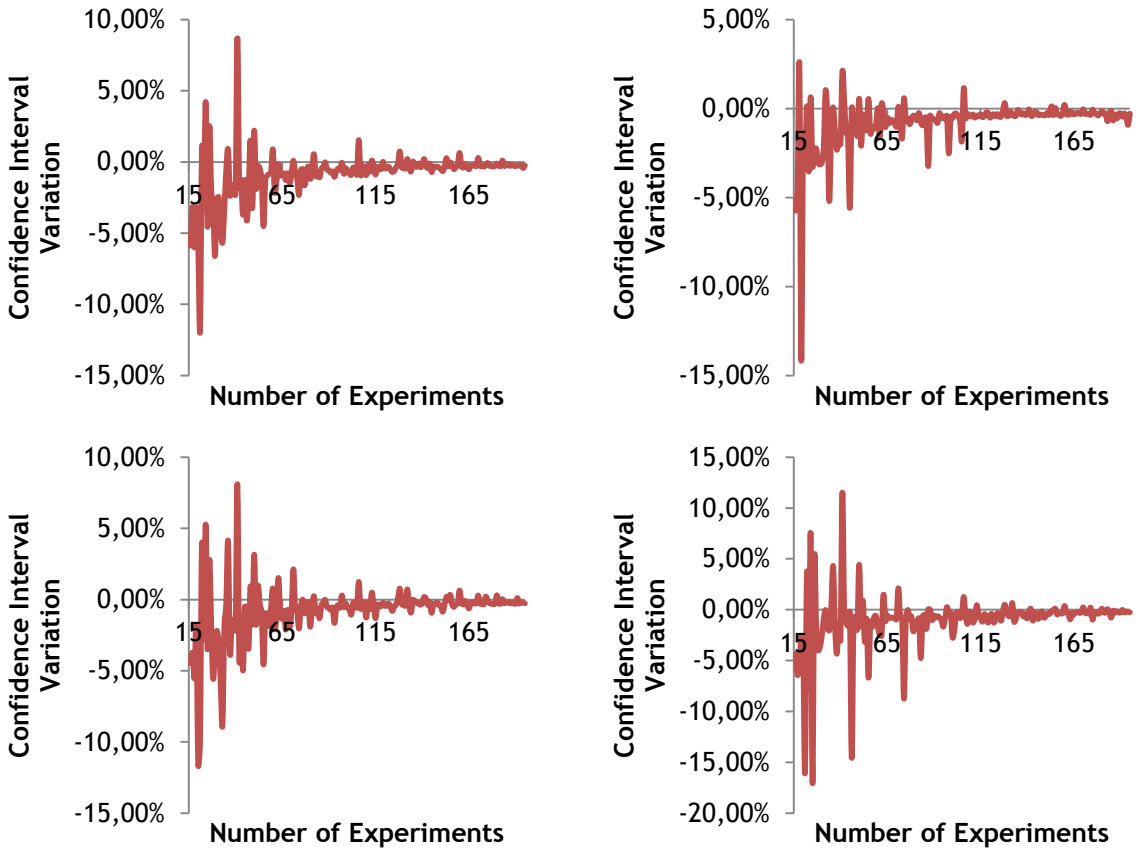


Figure 24 - Confidence Interval bounds variation with each experiment performed for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

The Confidence Interval variation gets lower than 10% after experiment 46. This information confirms the lower robustness of this estimation.

6.1.3 Model parameters using I-optimal

The third considered parameters were the parameters obtained in section 4.2.2 but using an I-optimal method to determine the next experiment. These parameters are shown in table 7.

Table 7 - Initial Estimation Value

Parameter	Estimated value
k_0	5.519E-01
E_A	2.000E+05
K_{MCH}	4.203E+00
K_{Tol}	1.358E+00

Figures 25 to 28 shows the parameter estimation, the confidence intervals and the t-value.

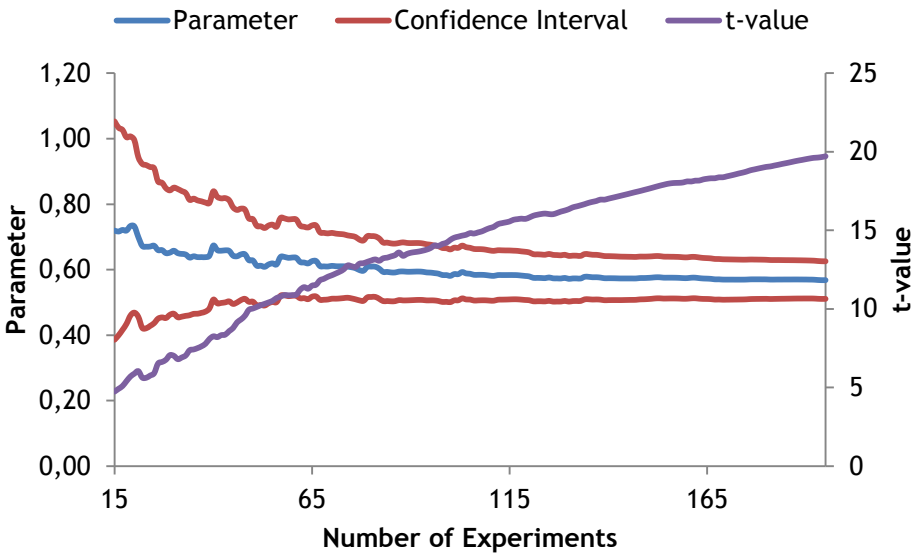


Figure 25 - k_0 Parameter estimation, confidence interval and t-value

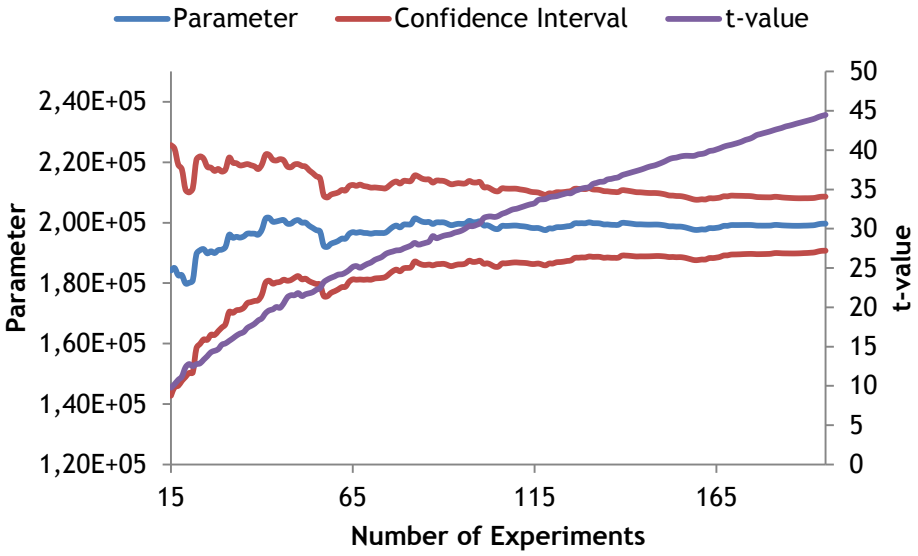


Figure 26 - E_A Parameter estimation, confidence interval and t-value

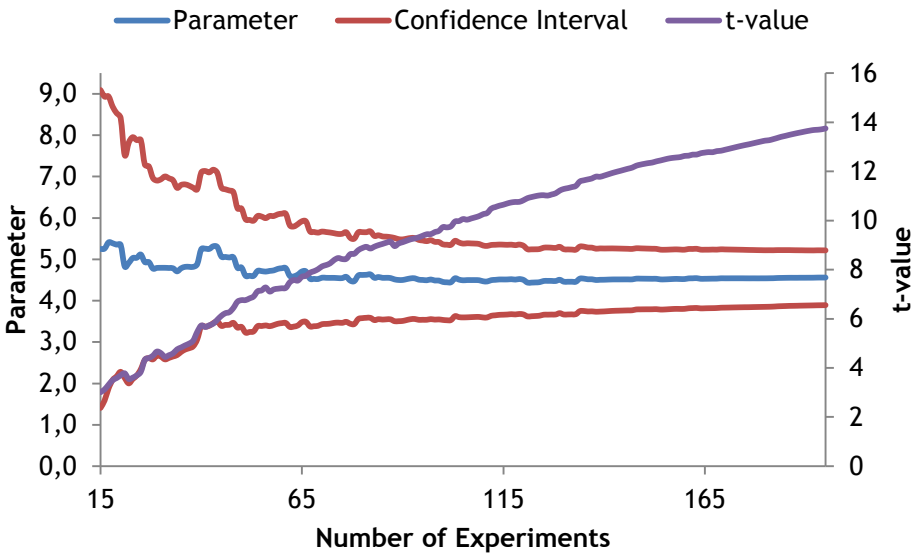


Figure 27 - K_{MCH} Parameter estimation, confidence interval and t-value

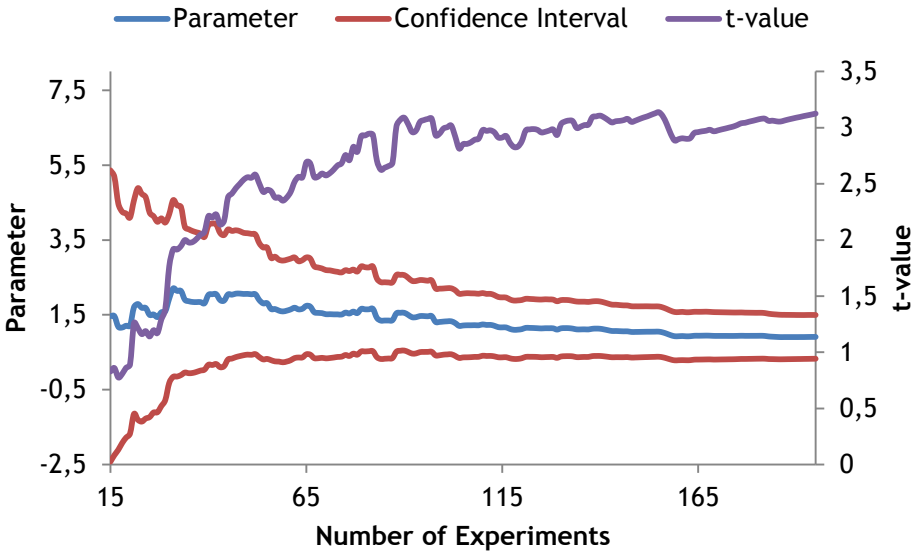


Figure 28 - K_{Tol} Parameter estimation, confidence interval and t-value

From the analysis of these figures, it is possible to conclude that the I-optimal method is less robust than the D-optimal one.

For an easier analysis of the confidence interval, the plot of the difference between the upper and lower bound was drawn. That plot is represented in figure 29.

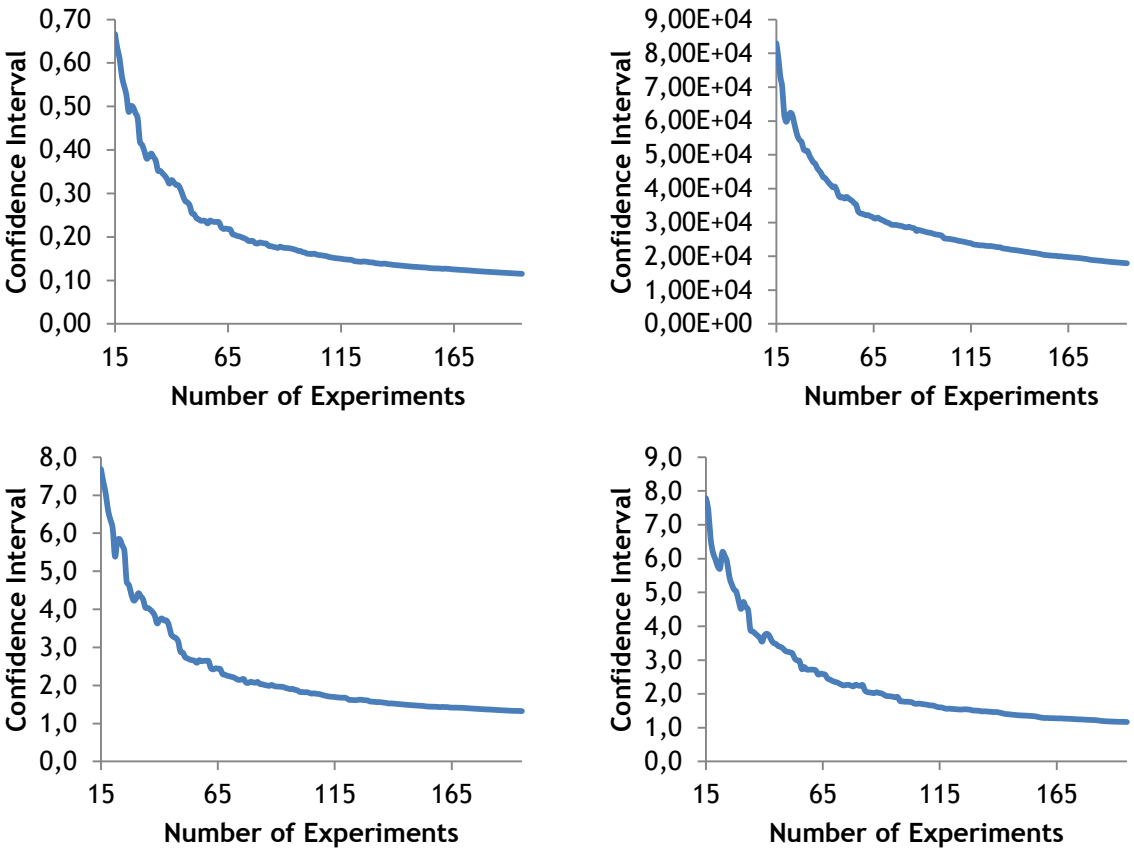


Figure 29 - Confidence Interval Bounds difference for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

As in the previous case, it is possible to verify the higher instability of the confidence intervals in this figure.

The variation of the parameters calculated after each experiment is represented in figure 30.

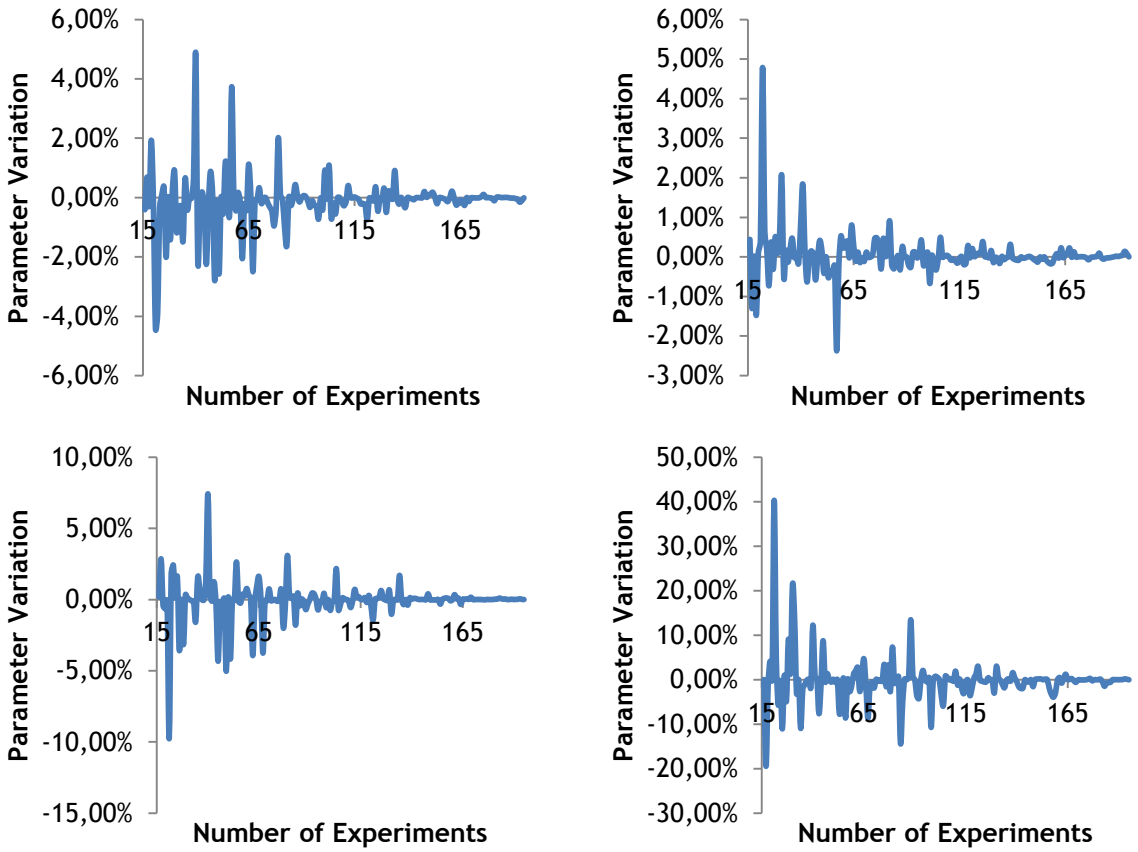


Figure 30 - Parameter variation with each experiment performed for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

In this case, K_{Tol} only varies less than 10% after experiment number 98.

The I-optimal method is worse than the D-optimal one but, in this case, the parameters stabilize faster than considering the previous parameter estimation to calculate the Jacobian matrix.

The variation of the Confidence Intervals bounds after each experiment is represented in figure 31.

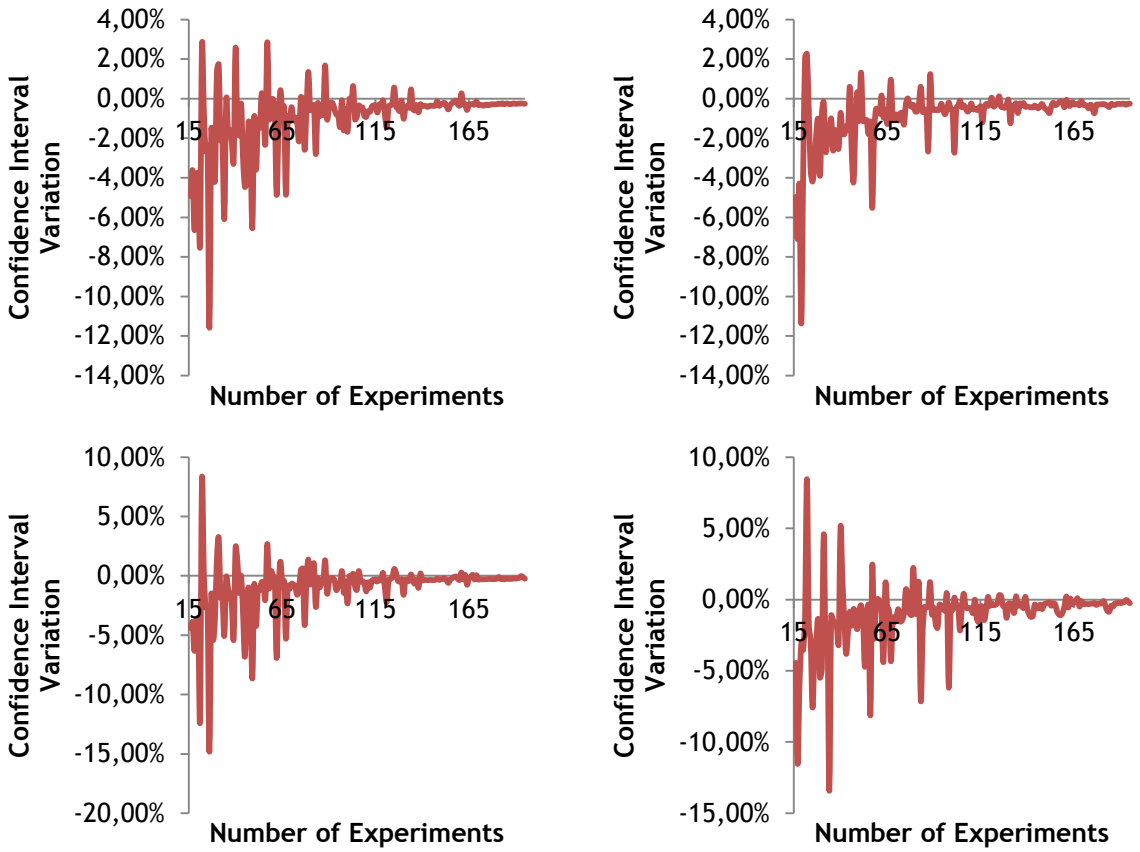


Figure 31 - Confidence Interval bounds variation with each experiment performed for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

The Confidence Interval variation gets lower than 10% after experiment 34.

6.2 Experimental Design Recalculating the Jacobian Matrix using the Calculated Parameter

In this subsection the Jacobian Matrix was calculated after every determined experiment.

To determine the next experiment, an I-optimal method was used. The initial estimation parameters are shown in table 8.

Table 8 - Initial Estimation Value

Parameter	Estimated value
k_0	1.000E+00
E_A	1.000E+05
K_{MCH}	1.000E+00
K_{Tol}	1.000E+00

Figures 32 to 35 show the parameter estimation, the confidence intervals and the t-value.

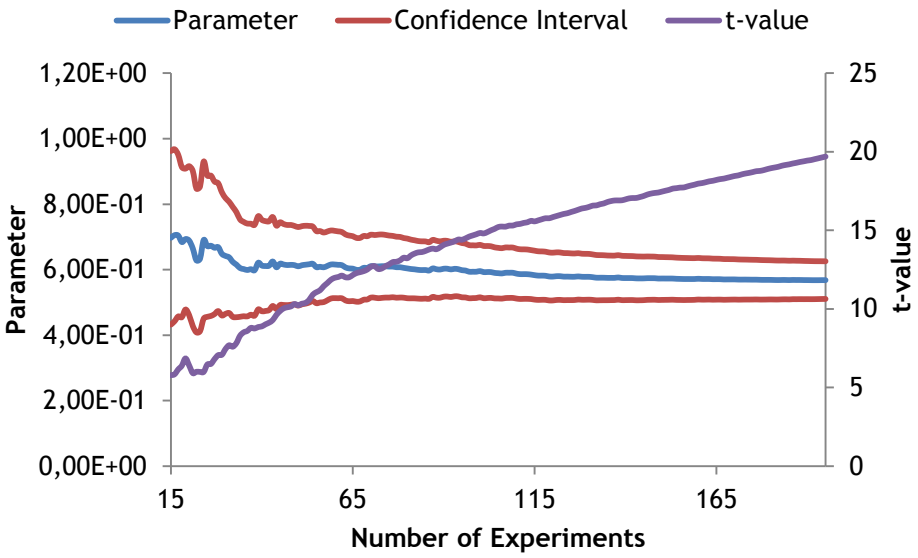


Figure 32 - k_0 Parameter estimation, confidence interval and t-value

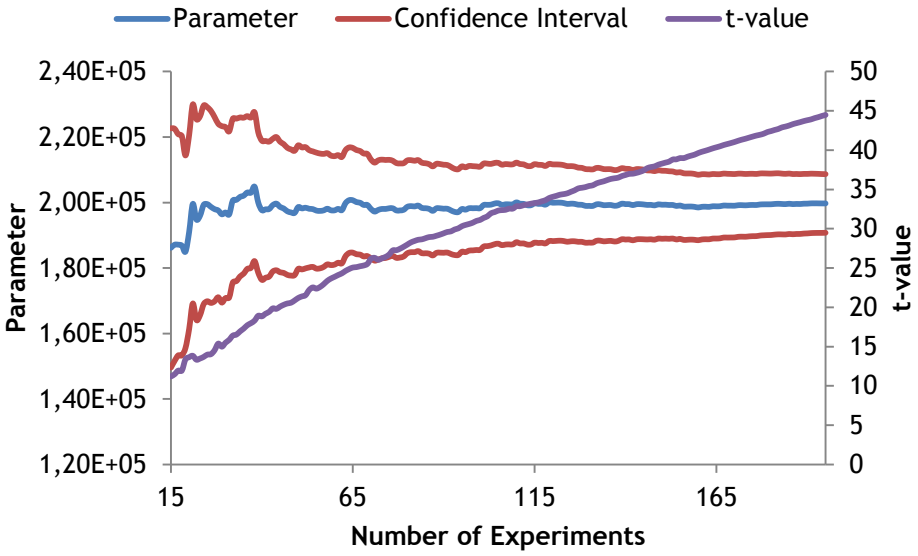


Figure 33 - E_A Parameter estimation, confidence interval and t-value

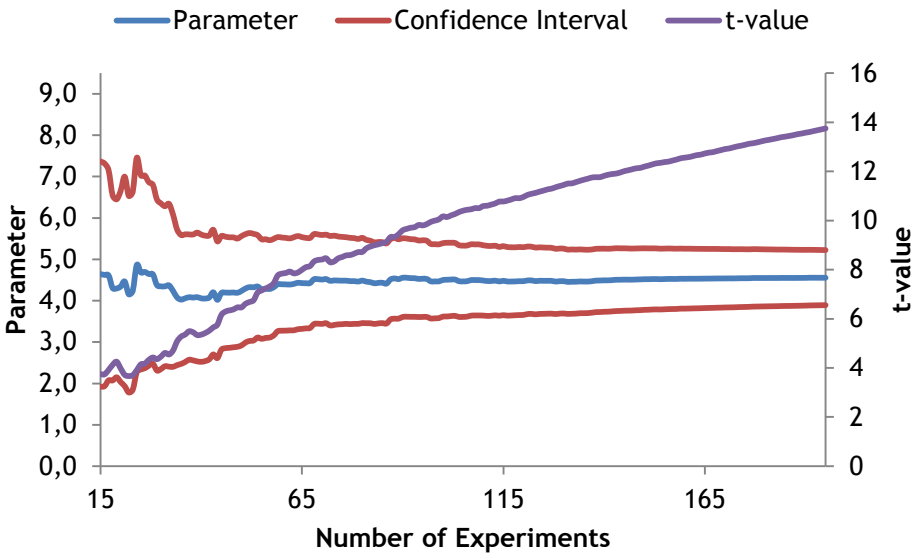


Figure 34 - K_{MCH} Parameter estimation, confidence interval and t-value

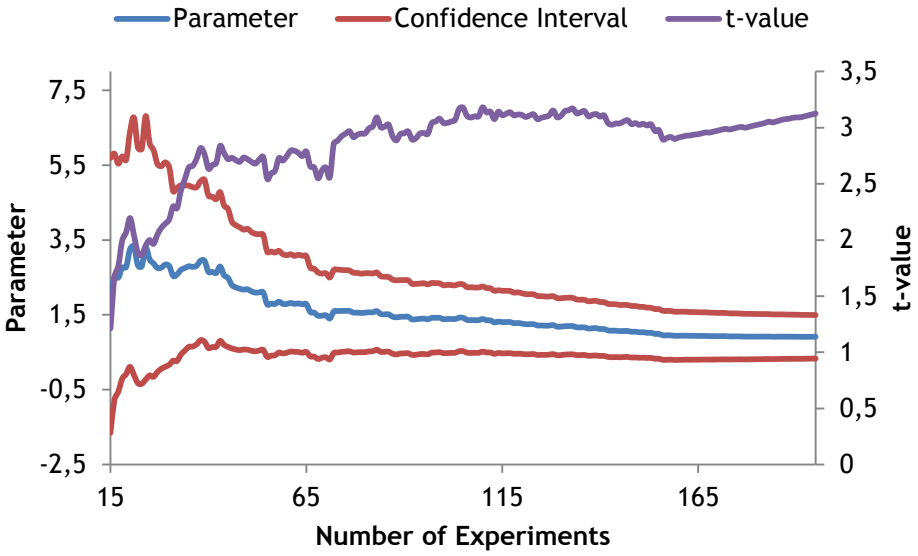


Figure 35 - K_{Tol} Parameter estimation, confidence interval and t-value

Analyzing these figures it is possible to see that in the first experiments exists a high instability but that instability is reduced really fast.

For an easier analysis of the confidence interval, the plot of the difference between the upper and lower bound was drawn. That plot is represented in figure 36.

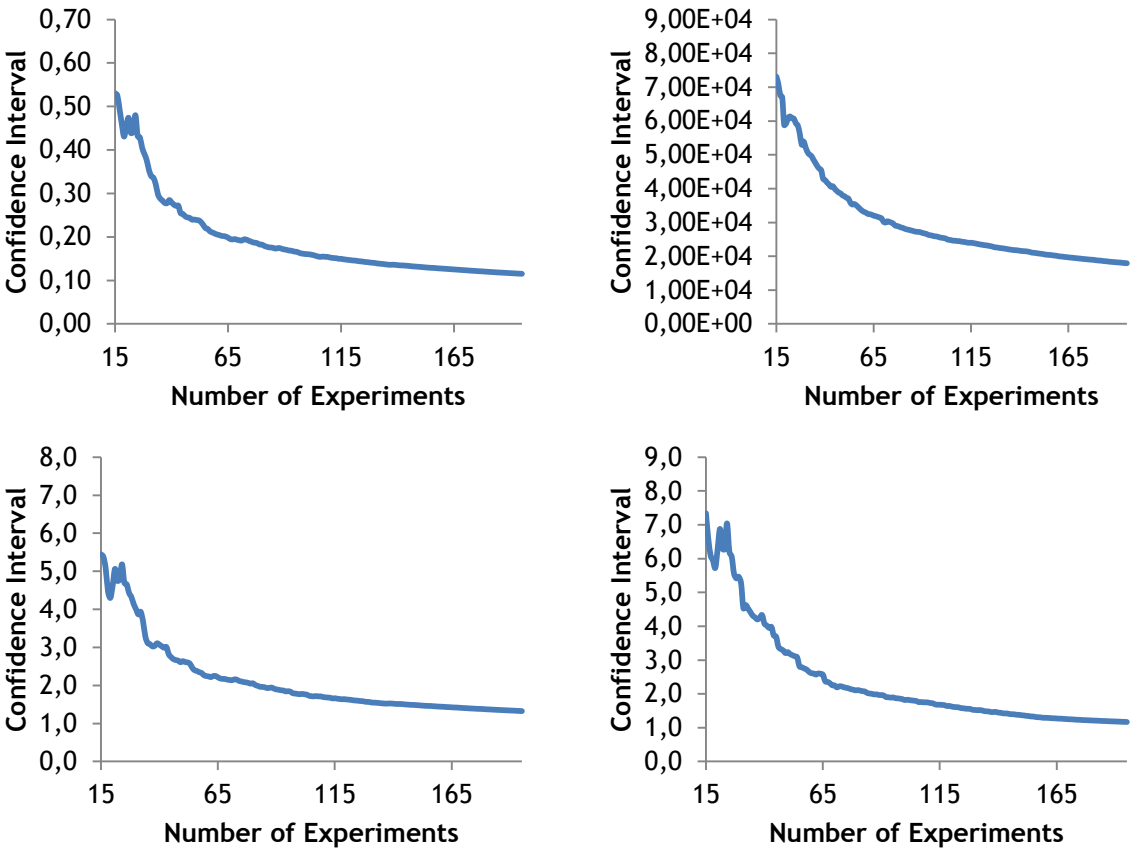


Figure 36 - Confidence Interval Bounds difference for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

The same instability is possible to be seen in the Confidence Intervals.

The variation of the parameters calculated after each experiment is represented in figure 37.

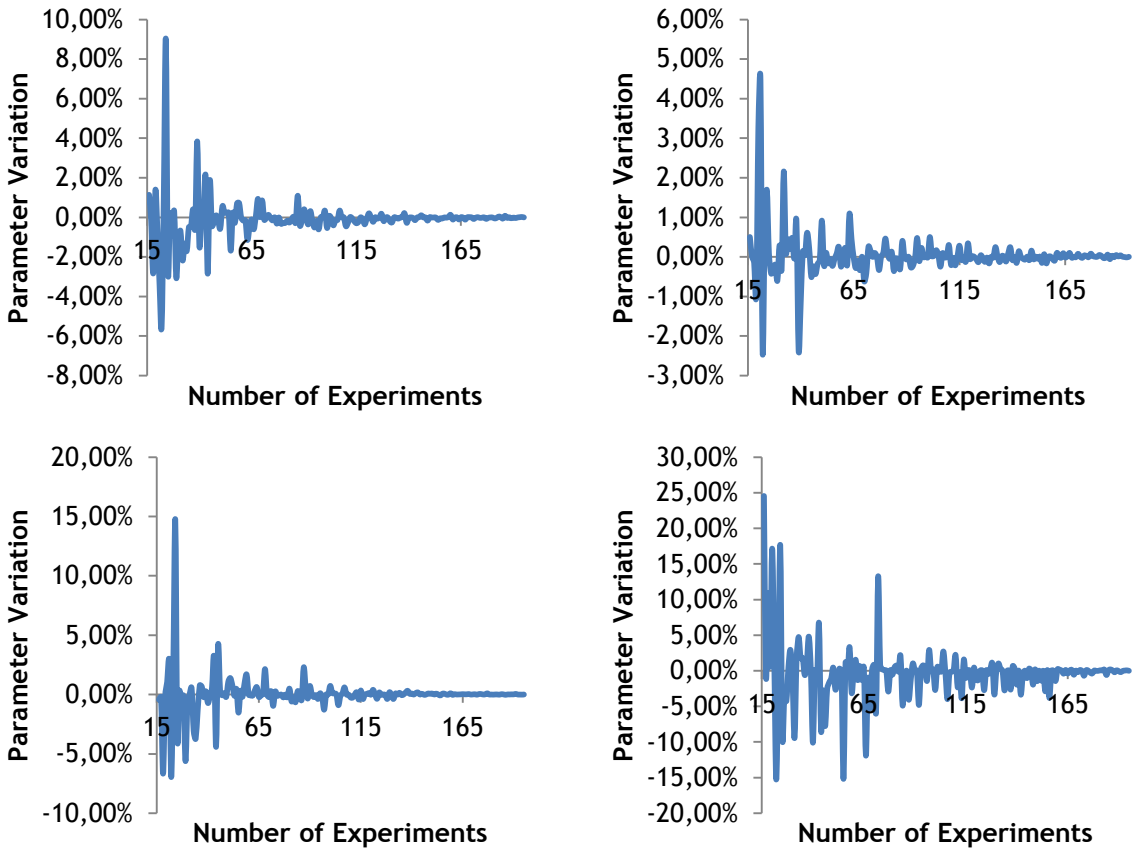


Figure 37 - Parameter variation with each experiment performed for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

In this situation, K_{Tol} only varies less than 10% after experiment number 72.

In this case, we have the variation being lower than 10% in the lowest experience number. That shows that, despite the earlier instabilities, the method starts to stabilize before all the methods considered before.

The variation of the Confidence Intervals bounds after each experiment is represented in figure 38.

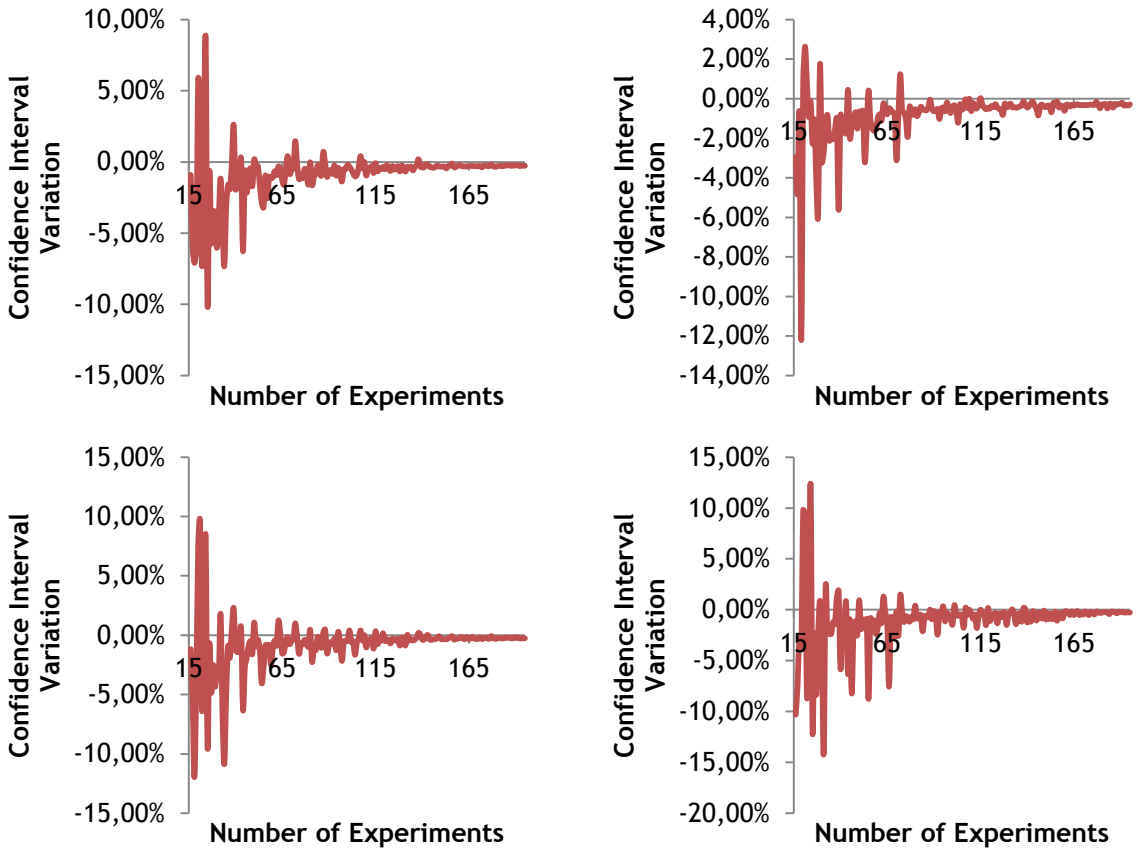


Figure 38 - Confidence Interval bounds variation with each experiment performed for a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

The variation of the confidence interval is lower than 10% after 34 experiments. This is lower than in the first method. This difference is not much significant and one of the causes can be that as we are using less fitting parameters to all the dataset in the beginning, it determines some experiments that would not be as good as in the first method, but after a low number of experiments this is corrected.

6.3 Conclusions

From subsection 6.1, it is possible to conclude that using a D-optimal design method is preferable comparing to the I-optimal one, since the parameter estimation and the confidence intervals start to have a variation lower than 10% earlier.

Comparing different parameters estimations that have a physical meaning, the best estimation would be to use the true parameters to calculate the Jacobian matrix, with the objective to determine the next best experiment. As the true parameters usually are unknown this case is somehow impossible to perform when trying to implement this method in kinetic models determination.

From section 6.2, it is possible to conclude that, despite the earlier instability in the method that recalculates the Jacobian matrix after each experiment, this method is the one in which the variation of the parameters is less than 10% earlier.

From the experimental point of view, this method is the best considered method to be implemented since it uses a plausible estimation of the parameters and takes into account the performed experiments, as well as the obtained parameters, to determine the next best experiment.

7 Conclusion

From Section 3, it can be concluded that the objective function that best optimizes our experimental data is the sum of squares of absolute deviations. This also comes from the consideration that the industrial units operate at high conversion, and that the higher conversions need to be accurately predicted.

From Section 4, it was shown that the power law model that considers the inverse reaction, despite being a good approximation of our reaction kinetics, does not describe correctly the inherent physics of the problem. The model that best describes the experimental data and the physics is a Single site surface reaction according to the Eley-Rideal mechanism, with the generation of the second hydrogen molecule as controlling step.

From the results of Section 5, it is revealed that the noise that will has the lowest impact on the results in terms of confidence interval and parameter estimation is the random relative noise. Both the fixed relative noise and the random relative Gaussian noise are found to be almost equal and inferior to the random relative noise.

Finally, from Section 6, one can conclude that using a D-optimal design method is preferable than using an I-optimal one. The best initial parameter estimates to calculate the Jacobian matrix, which is used to calculate the experimental sequence, are the true parameters. The method that recalculates the Jacobian matrix after each experiment, despite having some instability associated with the first experiments, is the best method to be used in the experimental sequence determination. With this experimental design method, it is possible to use less than eighty experiments, thereby reducing the number of experiments to be performed to less than half.

7.1 Accomplished objectives

The accomplished objectives were:

- The test of four different kinetic models and the determination of which of them best represents the experimental data;
- The analysis of the evolution of the parameters precision with different objective functions and noise;
- The experimental design to determine the number of needed experiments to obtain a certain parameters precision.

7.2 Limitations and Future Work

This work was performed using a single and simple reaction to optimize the number of experiments.

Future work should be the performance of the determination of experiments for a multiple response system in order to determine the next best experiment to perform with different models, so as to determine if it is possible to use this method in research labs.

7.3 Final Appreciation

I personally find the topic very interesting because it holds a lot of possible benefits for the future model developments. Especially the sequential design approach can provide a very significant contribution since it may allow to substantially reduce the number of experiments that are required to derive a kinetic model.

References

1 Verstraete J. (1997) Kinetische studie van de katalytische reforming van nafta over een Pt-Sn/Al_2O_3 katalysator, Universiteit Gent.

2 Fogler H.S. (2011) *Essentials of Chemical Reaction Engineering*, Prentice Hall. ISBN: 9780137146123.

3 Gates B.C., Katzer J.R., Schuit G.C.A. (1979) *Chemistry of catalytic processes*, McGraw-Hill, New York. ISBN: 9780070229877.

4 Edgar T.F., Himmelblau D.M., Lasdon L.S. (2001) *Optimization of chemical processes*, McGraw-Hill. ISBN: 9780070393592.

5 Verstraete J. (2012) Cinétiques réactionnelles - Chapitre 6 : Analyse cinétique -, 2012.

6 Jones B., Goos P. (2012) I-optimal versus D-optimal split-plot response surface designs. Available at: https://ideas.repec.org/p/ant/wpaper/2012002.html.

7 Wheeler R.E. (2004) opt Federov. AlgDesign. Available at: http://www.r-project.org/.

Appendix A. Four-Parameter Power Law

The other power law considered was the one given by equation B.1.

$$r = k_0 \times e^{\frac{E_A}{R}\times\left(\frac{1}{T}-\frac{1}{T_0}\right)} \times {p_{MCH}}^{n_1} \times {p_{H_2}}^{n_2} \quad \text{(B.1).}$$

The obtained parameters and the corresponding confidence intervals and t-values are presented in Table 9.

Table 9 - Parameter estimation and confidence interval for a Power Law with 4 parameters

Nr.	Parameter. Estimation	Lower Limit	Upper Limit	t-value
k_0	1.232E-01	7.965E-02	1.667E-01	5.661
E_A	1.791E+05	1.665E+05	1.916E+05	28.54
n_1	3.987E-01	3.474E-01	4.500E-01	15.54
n_2	-1.090E+00	-1.225E+00	-9.552E-01	-16.14

The n_2 estimation is negative because the Hydrogen adsorption favors the reverse reaction that was not taken into account by this mechanism.

The parity plot for this model is shown in figure39.

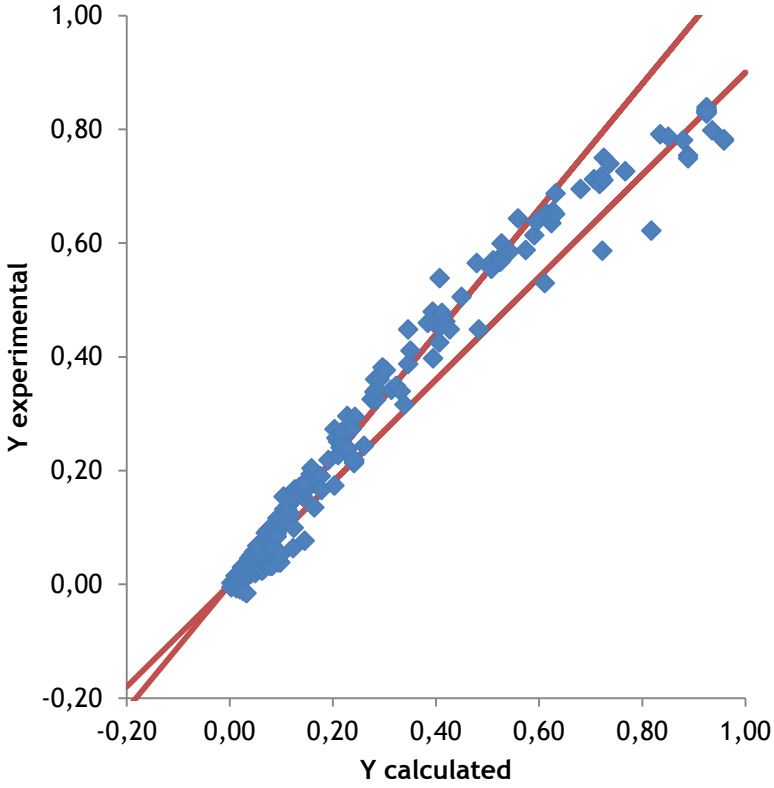


Figure 39 - Parity plot considering an experimental conversion relative error of 0.1 for a Power Law with 4 parameters

It is visible that the representation of the experimental and calculated conversions does not follow a linear regression along the bisector, and consequently, it is possible to conclude that this kinetic model does not correctly represent the experimental data.

The F-value obtained for this model was 3108.356

Appendix B. Effect of Noise on Parameter Values and t-Values

B.1 Random Relative Noise

B.1.1 Fixed signals ratio

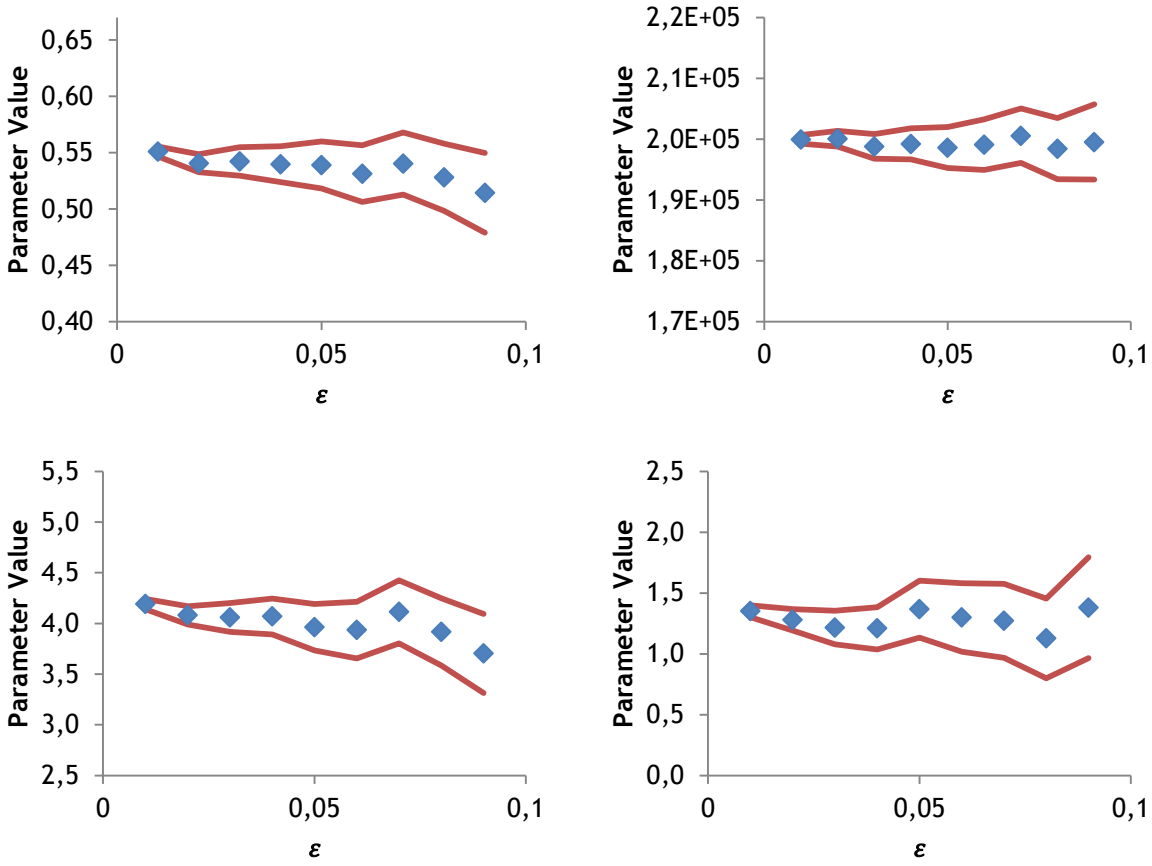


Figure 40 - Parameters estimation for the different random noise values with fixed signals ratio, where a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

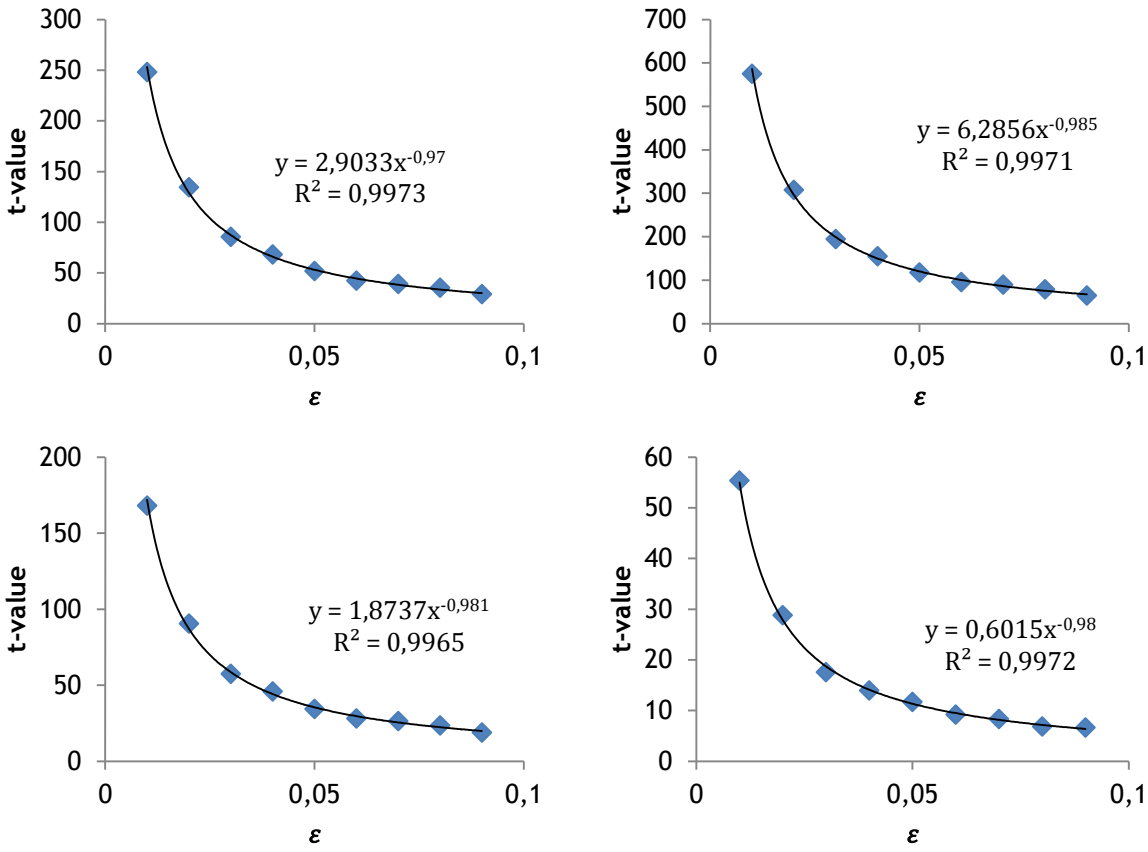


Figure 41 - t-value for the different random noise values with a fixed signal ratio where a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

B.1.2 Random signals ratio

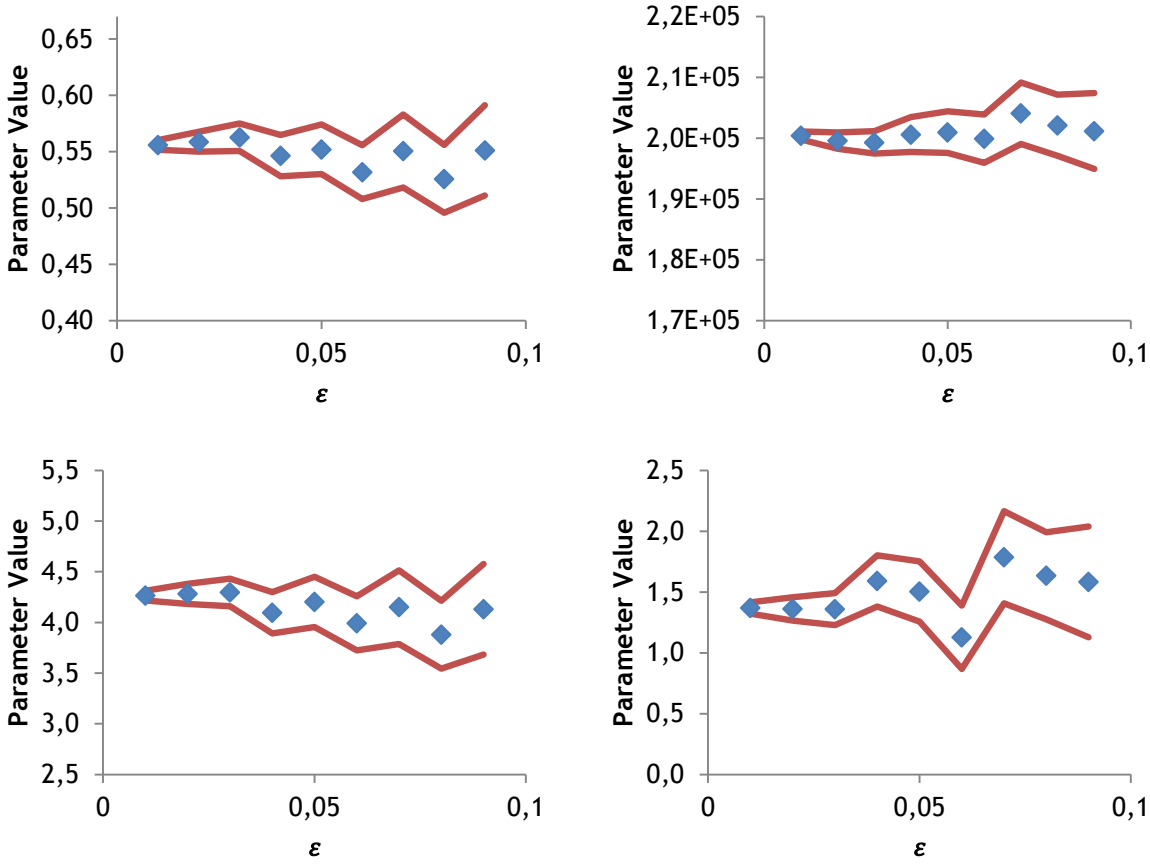


Figure 42 - Parameters estimation for the different random noise values with random signals ratio, where a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

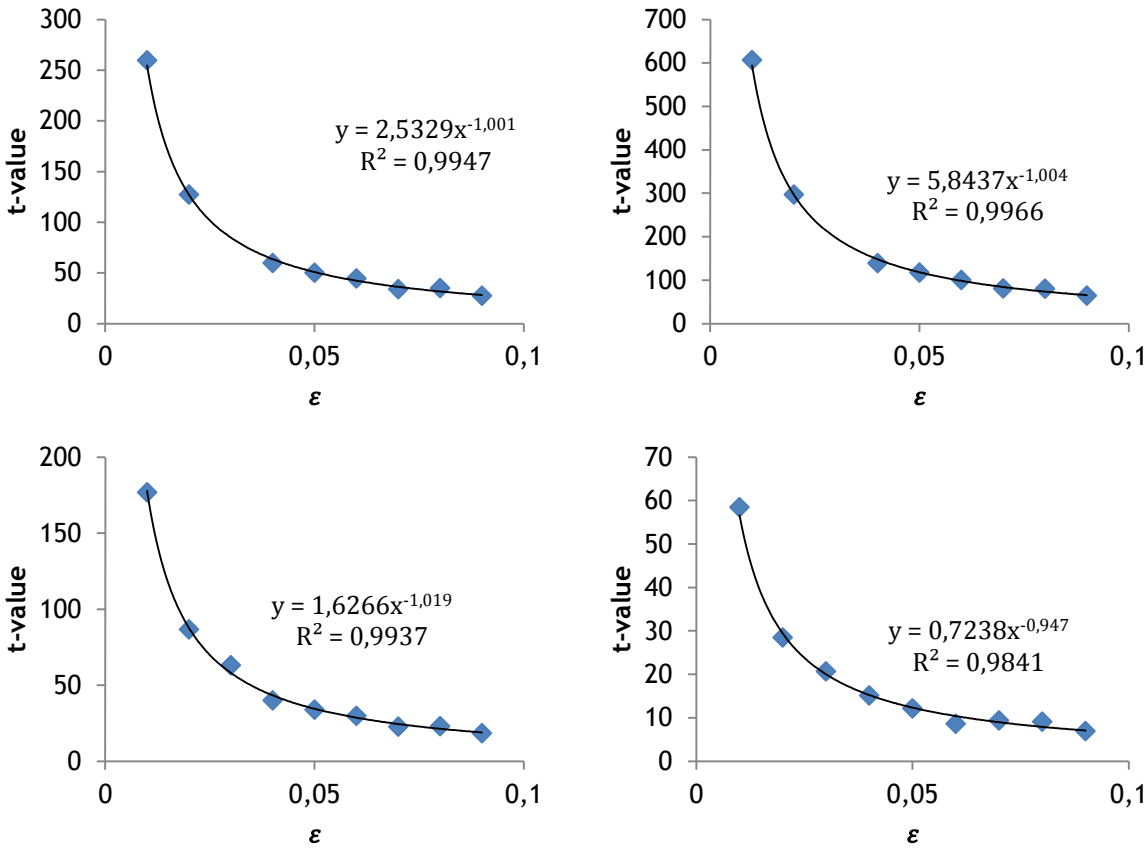


Figure 43 - t-value for the different random noise values with a random signal ratio where a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

B.2 Relative Gaussian Noise

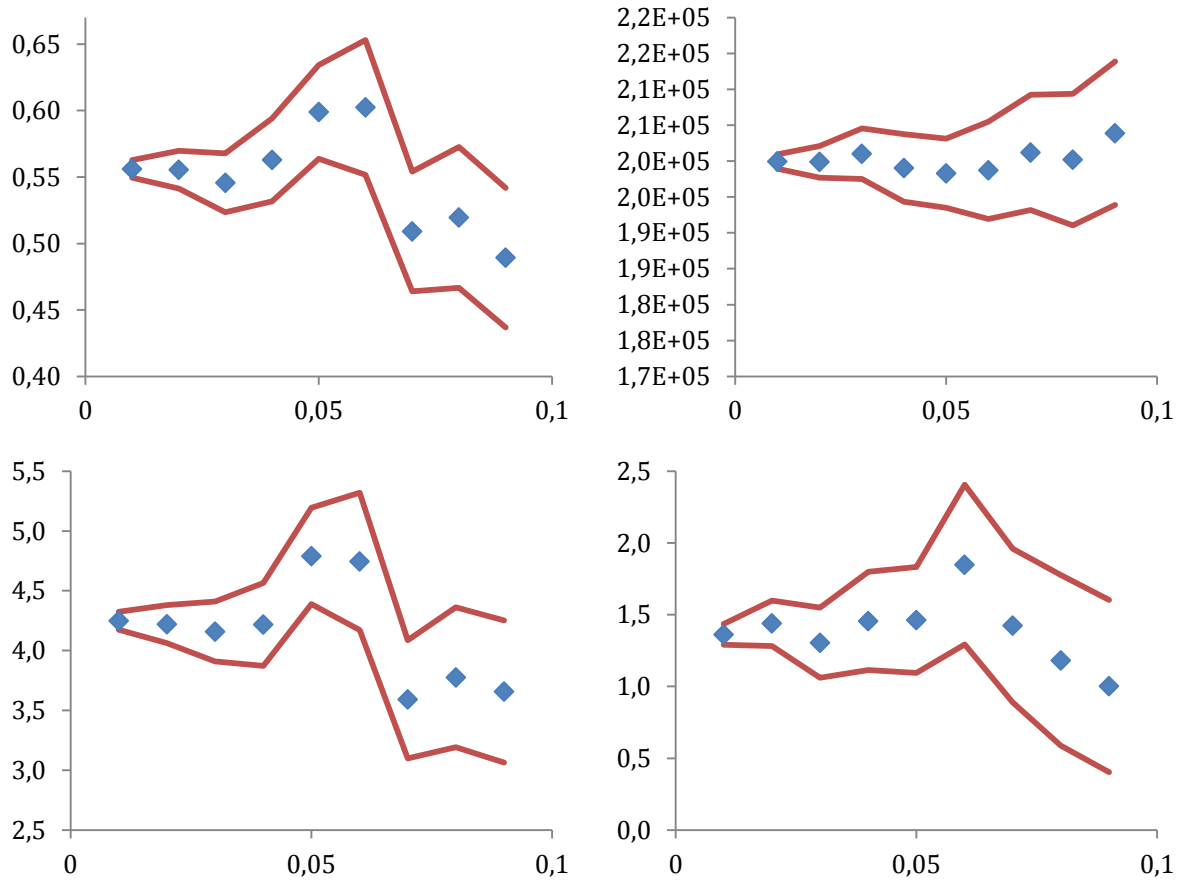


Figure 44 - Parameters estimation for the different Gaussian noise values where a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}

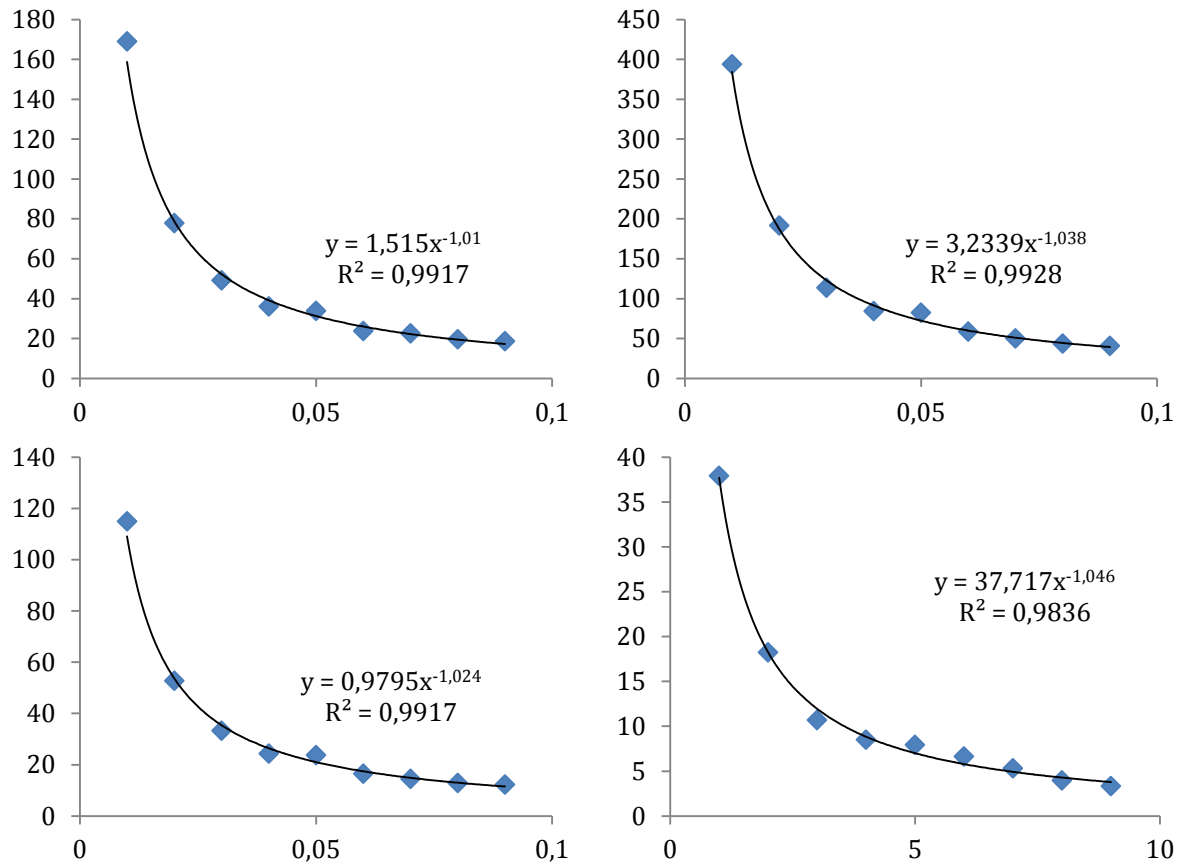


Figure 45 - t-value for the different Gaussian noise values in which a) k_0, b) E_A, c) K_{MCH} d) K_{Tol}