Otto-von-Guericke-Universität Magdeburg

Fakultät für Verfahrens- und Systemtechnik

Model-based stochastic optimization of palm oil deodorization via Short-Path-Distillation towards 3-MCPD fatty acid ester formation and major oil quality parameters

Dissertation

zur Erlangung des akademischen Grades

Doktoringenieur (Dr.-Ing.)

von Dipl.-Ing. Tim Christoph Rudolph

geboren am 17. Juli 1976 in Essen

genehmigt durch die Fakultät für Verfahrens- und Systemtechnik

der Otto-von-Guericke-Universität Magdeburg

Promotionskommission: Prof. Dr.-Ing. Dominique Thévenin (Vorsitz)

                       Prof. Dr.-Ing. habil. Dr. h. c. Lothar Mörl (Gutachter)

                       apl. Prof. Dr. Waltraud Kahle (Gutachterin)

                       Prof. Dr. Gerald Warnecke (Gutachter)

Eingereicht am: 27. Juni 2016

Promotionskolloquium am: 14. Dezember 2016

Danksagung

Abstract

Due its superior technological properties in food manufacturing, palm oil represents 75% of the worldwide edible plant oil production, which is one of the most important sectors of the global food industry. For the deodorization process, a distillative sub-process of crude oil refining to remove undesired volatile minor compounds, the heat-induced formation of 3-monochloropropane-1,2-diol and glycidol, as well as their fatty acid ester, could be proven. These substances have been defined as being potentially carcinogenic, tumorigenic and nephrotoxic to humans by regulatory authorities and research groups, resulting in the definition of threshold levels for refined oils and manufactured foods. Recent studies, aiming at minimization of these contaminants, propose the Short-Path-Distillation (SPD) as promising alternative for the standard deodorization process due to comparatively mild thermal distillation conditions at high vacuum and other features. In this regard, the SPD is systematically analyzed and optimized model-based by an innovative stochastic modelling approach in this work. This approach comprises the Response Surface Methodology (RSM) for model derivation and Bayesian model analysis. The RSM provides a set of linear process models which comprise the effects and interactions of SPD process variables on the palm oil quality parameters. These models are just valid for the individual process analyzed but provide informations of general validity on the qualitative nature of effects and interactions that aid design and scale-up of comparable SPD processes. Subsequently, a coherent Bayesian model analysis methodology is implemented and applied to the set of linear models to account for the uncertainty of models and results with respect to the error of experimentation. By that approach, probability distributions over the range of alternative results are estimated that provide the mode estimate, expected value estimate and corresponding uncertainty measures (standard deviation), which are applied to the analysis of the SPD process. The Bayesian estimation of parameter values of the linear model terms (identified by RSM) is achieved by estimation of the Bayesian posterior parameter probability distribution via Gibbs-sampling (special case of the Metropolis-Hastings-Algorithm (Markov-Chain-Monte-Carlo method)). This sampling process is accelerated by an integrated sub-algorithm as proposed by Müller (1991). Alternative models for a defined quality parameter (RSM provides models of 1st and 2nd grade) are discriminated by estimation of the Bayesian relative model probability according to the method proposed by Chib-Jeliazkov (2001). The Bayesian validation of models, as well as Bayesian evaluation of model performance, model predictions and experimental results, is achieved by implementation of the method proposed by Geweke (2007). To ensure a correct implementation and application of all stochastic methods and algorithms for the real-life problem of SPD optimization, a validation of method implementation and understanding is achieved by initial application to a simple model and simulated experimental data (derived from that model by random addition of noise). All experiments achieved for model derivation show concentrations of 3-MCPD and related substances (such as glycidol) that are negligible and / or within the error experimentation. Thus, the derivation of statistically significant model terms is not possible. As the measured values are also far below all discussed threshold levels it can be stated, that the SPD is definitely capable of meeting the corresponding process and product quality requirements. The RSM provides model proposals of 1st and 2nd grade for the oil quality parameters rancimat, acid value and tocopherol, which are further discriminated and analyzed by Bayesian means. In this regard, process settings are identified and experimentally verified, which ensure the compliance with oil quality standards with the desired stochastic certainty. With respect to method and product quality requirements, evidence is provided in this work for the outstanding performance of the SPD and the stochastic methods applied.

Content

Index of Tables

Index of Figures and Pictures

XIII

Index of Equations

Index of Abbreviations

JECFA	Joint Expert Committee on Food Additives
m	Number of factors / process variables in SDoE
M	Molecular weight; Bartlett-criterion; model; matrix of normalized Eigenvectors
MAG	Monoacylglycerine
MCMC	Markov-Chain-Monte-Carlo
M_{crit}	Critical value of Bartlett-criterion
MM-Model	Michaelis-Menten-Model
MRI	Max-Rubner-Institute
n	Number of experiments in a SDoE (single repetition)
n_T	Total number of experiments achieved (for multiple repetition of SDoE)
N_0	Number of trials for the center point of a CCD
n_0	Number of trials for the center point of an orthogonal CCD
N_{tot}	Total number of experiments of a CCD
n_{tot}	Total number of experiments of an orthogonal CCD
N_{st}	Number of trials for the starpoints of a CCD
n_{st}	Number of trials for the starpoints of an orthogonal CCD
$n_{Stirrer}$	Stirrer rotation speed
N_w	Number of trials for the core of a CCD
n_w	Number of trials for the core of an orthogonal CCD
p	Pressure; number of parameters; probability
P_{Pump}	Pump power (frequency)
$\widehat{P}$	Posterior probability estimate
q	Probability of proposal
QS	Summed squares
R	Gas constant
RSM	Response Surface Methodology
SCF	Scientific Committee on Food
SDoE	Statistical Design of Experiments
SPD	Short Path Distillation
ST	Star points of the CCD
T	Temperature; matrix transpose
TAG	Triacylglycerine
$T_{Condenser}$	Temperature of condenser
TDI	Tolerable Daily Intake
$T_{Evaporator}$	Temperature of evaporator
v	Relative variance
W	Core of CCD
WHO	World Health Organization of the United Nations
WW	Interaction of process variables
x	Variable of modell equation
x_{max}	Maximum of analysis range of variable x in SDoE

x_{min}	Minimum of analysis range of variable x in SDoE
$\tilde{x}$	Normalized variable of model equation
$\bar{x}$	Mean value of process variable range
x^*	Normalized extremum of model equation
y	Experimental result for target value; exp. data vector
$\hat{y}$	Simulation result for target value
$\bar{y}$	Arithmetic mean for target value at defined process setting
z	Function of y

1 Introduction

The production of edible oils and fats represents one of the most important sectors of the food industry worldwide due to its quantitative and functional importance for the production of many foods. In 2003, the total global production of edible oils and fats reached approximately 105 million tons, of which 75% were applied for human consumption. Among edible plant oils, palm oil represents the most significant type, followed by soybean, rapeseed and sunflower oil. From 1999 to 2009, the global production of crude palm oil (CPO) nearly doubled by increasing from 20.6 to 45.4 million tons. Due to its superior oxidation and heat stability, as well as the fact that it maintains a solid consistency at room temperature, palm oil is of special importance for the food industry, particularly for bakery products, drippings, margarine and confectionaries (OVID, 2015; USDA, 2010).

Aside from its technological advantages, palm oil represents a valuable source of nutrients as it provides many liposoluble vitamins, such as tocopherols and carotenoids, essential polyunsaturated fatty acids and secondary phytonutrients (Rimbach et al., 2010; O'Brien, 2009 and BVE, 2010). Most crude oils are not suitable for human consumption and food production due to the fact that various associated minor compounds, such as free fatty acids (FFA), oxidation products, phospholipids, carotenoids and pesticides, negatively affect the sensory, nutritional and technological product properties.

It has to be mentioned here that, besides its economical, technological and nutritional benefits, the environmental, health and ethical aspects of modern palm oil production have been recently discussed very critically. A further detailed analysis and discussion of this issue is not objective of this work.

1.1 Modified processes for control of 3-MCPD fatty acid ester formation

Thus, a preceding treatment of the oil by a sequence of different sub-processes, known as the refining process, is necessary to remove these undesired substances. One of these sub-processes is the so-called deodorization, which is a distillation process that applies an additional steam strip to remove volatile compounds at temperatures of up to 250°C (Bockisch, M., 1993, pp. 484-486; 531-537). In 2006, the heat induced contaminants 3-monochloropropane-1,2-diol (3-MCPD) and glycidol, as well as their fatty acid esters (3-MCPD-FE and G-FE), could be detected for the first time in refined fats and oils (Zelinková et al., 2006, pp. 1290-1298). The detection of these contaminants was confirmed by CVUA Stuttgart (Weißhaar, R., 2008). For free 3-MCPD nephrotoxic and tumorigenic properties could be discovered at higher doses for rats (Sunahara et al., 1993).

Due to these facts, the Scientific Committee on Food (SCF) of the EU-Commission and a FAO/WHO committee (JECFA) defined a tolerable daily intake (TDI) of 2 µg·(kg body

weight)$^{-1}$ of 3-MCPD by taking a safety factor of 500 into account. Recently, the International Agency for Research on Cancer (IARC) defined free 3-MCPD as being "potentially carcinogenic to humans". In the course of developing suitable analytic methods for 3-MCPD-FE detection, also high amounts of G-FE, which is also considered as being harmful to human health, were detected especially in refined palm oil (Bakhiya et al., 2011; Habermeyer et al., 2011). Free glycidol exhibits mutagenic and carcinogenic properties and has been defined as being "probably carcinogenic to humans" by the IARC.

Furthermore, the ALARA-principle is valid for G-FE, which means that its content in refined edible plant oils and manufactured products has to be kept "as low as reasonably possible". The German Federal Institute of Risk Assessment (BfR) assumes that 3-MCPD-FE and G-FE are enzymatically cleavaged into free 3-MCPD / G and free FFA in the human gastrointestinal tract by 100%, which means that all threshold levels and recommendations that have been defined for the free substances so far are also valid for the ester. The European Food Safety Authority (EFSA) agrees with that assessment, which is additionally supported by recent research on the metabolism of 3-MCPD-FE and G-FE (Creuzenberg and Berger-Preiß, 2011; BfR, 2009). Various studies have provided evidence that the deodorization process, which is a fundamental part of plant oil refining, is crucial for the formation of 3-MCPD-FE and G-FE due to the exposure of the oil to high temperatures (Hrncirik, 2010; Lin et al., 2010; Bhaggan / Werleman, 2010).

Consequently, many German and international institutes and authorities have defined specific threshold levels and initiated research projects aiming at the development of strategies to minimize the 3-MCPD-FE and G-FE content in refined plant oils. In a recent project initiated and funded by the Research Association of the German Food Industry (FEI), various minimization strategies were identified on laboratory scale which are capable of controling the content of these contaminants in the oil by modification of the refining conditions (Pudel et al., 2011) and a subsequent removal from the product by adsorbents (Strijowski et al., 2011). One promising strategy regarding the modification of the refining process is the replacement of the standard strip steam deodorization by a Short-Path-Distillation (SPD) process that is capable of significantly reducing the formation of heat induced contaminants by applying comparatively mild thermal conditions at high vacuum (Pudel et al., 2011; FEI, 2011, pp. 47-73; TTZ OvGU, 2012; ASW, 2012).

1.2 Short-Path-Distillation (SPD) and its potential

The potential of the SPD to minimize the formation of 3-MCPD-FE and G-FE, as well as to replace the standard deodorization process, is investigated in this work. For this purpose, the effects of the significant process variables of the SPD on the formation of these contaminants and major oil quality parameters are systematically analyzed for palm oil through the application of sophisticated stochastic methods of model-based process optimization. Palm

oil is chosen examplaryly as model oil due to its comparatively high 3-MCPD-FE and G-FE formation capacity and global economic importance. Process variable settings are identified and evaluated model-based at which a minimization of the formation of contaminants along with the compliance of the oil quality standards can be realized.

The deodorization of edible plant oils via SPD is a very complex process. Various physico-chemical effects and interactions influence the distillative separation of substances from the oil, some of which are supposed to be separated while others are supposed to remain in the product. The key effect in this context is the dependency of the vapour pressure of these substances on the temperature and ambient pressure which is given by the vapour pressure curve individually for each substance, thus enabling substance separation by differences in curve progression. Additionally, these effects and interactions depend on the complex technological design as well as various process variables to be adjusted. Thus, mechanistic modelling approaches are limited by the large number of effects, interactions and operative characteristics that should be considered in a suitable model, but unfortunately are not appropriately characterized or just too complex to be modeled properly in many cases.

1.3 Response Surface Methodology (RSM)

A novel process and product optimization approach is the modelling of processes by application of the Response Surface Methodology (RSM), which is a self-contained statistical modelling approach that is based on statistical design of experiments (DoE). The RSM identifies significant effects and interactions of the process variables (e.g. temperature and stirrer speed of SPD) towards the target values (e.g. oil quality parameter), which are additionally quantified by corresponding terms in a linear model that sums up these effects with respect to a certain target value. This approach is in line with the so-called "black box" principle which treats the outcome of a process (product quality parameters) as a function of the input (process variables) without considering the physico-chemical or technological sub-processes that convert the input into the output. The outcome of the RSM is actually a set of linear models of varying complexity that describe the dependency of a target value on the process variables being considered. The RSM therefore offers the opportunity to model and optimize a process without requiring a deeper knowledge on the fundamental physico-chemical sub-processes. This is based on a minimized number of statistically defined experimental process settings which maximize the information on the process contained in the corresponding experimental data. The lack of information regarding the key physico-chemical and technological processes inside the "black box" indicates that such a model is only valid for the operative and environmental process constants at which the experiments were achieved. Thus a model derived for a laboratory scale facility and process is only suitable for giving recommendations for other (e.g. industrial scale) comparable facilities and processes and for setting reasonable limits to the spectrum of process settings to be

examined for their optimization. For a detailed specific process modelling the experimental setup and model derivation has to be achieved separately for each diverging process setup (e.g. oil type, facility type and scale). The implementation of the RSM, as well as the analysis of its potential in process modelling and optimization, will be achieved by the laboratory scale approach of palm oil SPD that is presented in this work. Moreover a generalized algorithm is provided in this work that enables the derivation of experimental setup and process models for each diverging technological process of interest.

Through RSM, the parameter values of the potential model terms (effects and interactions on the target value respectively) are calculated directly from the experimental results with a subsequent statistical evaluation of significance. A major drawback of the RSM is, however, that the error of experimentation is not considered in terms of weighing and evaluation of experimental results and models / conclusions derived from the same. Furthermore, a comprehensive evaluation of reliability and uncertainty of the calculated results for model parameterization, model solutions, model performance and model selection is not provided. However, such an analysis and evaluation is inevitable for a self-contained model-based process optimization approach.

1.4 Bayesian model analysis

Bayesian analysis meets these requirements of a comprehensive model parameterization, discrimination and performance evaluation as it is an innovative and self-contained probabilistic approach that provides uncertainty measures for model parameterization, model selection and model solutions by means of probability distributions over the range of possible alternatives. These probability distributions are determined by updating the so-called 'prior distribution', which is based on previous knowledge (e.g. literature), by the information on the true underlying process contained in the experimental data. This updated version of the probability distribution is referred to as 'posterior distribution'. As the posterior probability of model parameterizations cannot be calculated directly, its estimation turns out to be the core problem and process of Bayesian analysis which is highly complex, both mathematically and computationally. Various approaches have been developed in an attempt to limit problems related to the complexity of this estimation. A sophisticated approach is the generation of a representative sample from the posterior distribution by application of Gibbs-sampling, a version of the Metropolis-Hastings-Algorithm which in turn is a Markov-Chain-Monte-Carlo approach (Chib and Greenberg 1995; Hastings, 1970; Tierney, 1994; Hartmann et al., 1974; Kreyszig, E., 1972; Chib, S. and Jeliazkov, I., 2001, pp. 270-281). As this algorithm requires significant computational effort Müller et al. (2004) proposed an efficient method to increase its speed and efficiency, which is referred to as Müller-algorithm in this work. This algorithm is incorporated into the Gibbs-sampling algorithm to accelerate Markov-Chain convergence towards the Bayesian posterior distribution. Additional problems are the discrimination of

alternative models by posterior model probability estimation, model validation and evaluation of model predictions.

Chib and Jeliazkov derived a self-contained and mathematically proven Bayesian random sampling method to estimate the posterior relative model probability as a model discrimination measure (Chib, S. and Jeliazkov, I., 2001, pp. 270-281). This method is based on the posterior parameter distribution sample and related data (e.g. likelihoods) provided by the Gibbs-sampling algorithm. Finally, a comprehensive evaluation of model performance requires model validation and analysis of model predictions. Unfortunately, Bayesian theory and application is less developed in this field. Geweke (2007) developed a coherent and mathematically proven theory for an exclusive Bayesian model validation and evaluation of simulation results in terms of probability. It is a random sampling method that is also based on the data output of the Gibbs-sampling. He further applied a case study to provide evidence for the practical suitability of this theory (Geweke, 2007).

In this paper, the Short-Path-Distillation of palm oil is modeled and optimized through the application of RSM and Bayesian analysis. Alternative models of varying complexity and with significant model terms are derived by RSM. The model terms are then parameterized via Bayesian determination of posterior parameter distributions and most probable parameters. The optimum process variable setting is then determined for each model. Additionally, the relative model probability (model discrimination), probability distribution of model predictions and model validity are determined by Bayesian means in an attempt to provide a comprehensive stochastic approach of model selection and evaluation of model performance. The model based calculated optimum process setting and, thus, model performance is experimentally proven in this document.

2 Theoretical background

In this chapter an introduction is given into the theoretical background in chemistry, process engineering and stochastic, wich is relevant for this thesis.

2.1 The 3-MCPD fatty acid ester and related substances in foods

This chapter describes the physico-chemical properties, formation and occurrence of 3-MCPD fatty acid ester and related substances, such as glycidol, in foods. The physiological impact and toxicity, regulative approaches and recent research efforts regarding these substances are also exposed in detail in this work.

2.1.1 3-MCPD

3-MCPD and glycidol are judged as so called food born toxicants (SGS, 2014). They are formed heat induced in thermal production processes of fatty and salty foods. Precursors are constituents of the raw material. A major process in that context is the conventional refining of edible oils and fats. Other critical products and processes will be described in the further subchapters. The following product groups are particulary pressured:

- Backery products
- Bread
- Drippings, chip fat and coconut oil
- Fish fingers
- Fried meat
- Hazelnut spreads
- Margarine
- Pizza
- Red seasoning sauce
- Refined baby nutrition
- Rissoles
- Sausages
- Smoked meat and fish products
- Spice paste

2.1.1.1 Physico-chemical characterization of 3-MCPD

3-monochloropropane-1,2-diol (3-MCPD) is part of the chloropropanol group and thus contains a polyvalent alcohol in its molecular backbone. In this molecular structure one or more hydroxyl groups of the alcohol are substituted by a chlorine atom. The chemical basis of 3-MCPD is glycerin, in which a terminal hydroxyl group is substituted by a chlorine atom (Hamlet et al., 2002b, pp. 619-631). A substance which is chemically closely related is 3-

MCPD-FE. In this fatty acid ester one or two hydroxyl groups of 3-MCPD are esterified with long chain fatty acids of variable length (see Figure 2.1). 3-MCPD-FE are nonpolar and consequently lipophilic. This is in contrast to the free form (polar), which is lipophobic (Baltes and Matissek, 2011, pp. 309–311; Hamlet et al., 2002b).
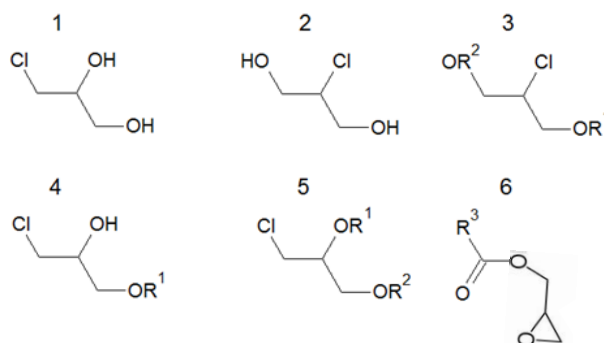
Figure 2.1: Alternative structures of 2-MCPD, 3-MCPD and glycidol (fatty acid esters respectively)
1): 3-MCPD; 2): 2-MCPD; 3): 2-MCPD-diester; 4): 3-MCPD-monoester;
5): 3-MCPD-diester; 6): Glycidyl-FE; $\mathbf{R^1, R^2}$ =Acyl residue of fatty acid; $\mathbf{R^3}$ =Alcyl residue of fatty acid

Fig 2.1 presents alternative structures of 2- and 3-MCPD in free form, 2- and 3-MCPD-FE and gylcidyl-FE (Crews et al., 2013).

2.1.1.2 Occurrence of 3-MCPD

Free 3-MCPD has been detected in doses that are critical to human health in refined edible oils and in foods that are produced from these raw materials. This chemical substance can be detected in cereal products, cheese and meat, as well as in artificial additives; a relevant example of an artificial additive is modified starch. Hamlet et al. (2002a/b) and Crews et al. 2002) thoroughly summarized foodstuffs containing 3-MCPD.

The formation of 3-MCPD begins as a heat induced thermal reaction. Therefore, its presence can be detected at higher contents in foods that have been subject to intensive heat treatment during the production process. Exemplarly the concentration of 3-MCPD increases from about 0,01 $mg \cdot kg^{-1}$ to 0,32 $mg \cdot kg^{-1}$ during the toasting of bread. Thereby, the content is strongly dependent on the browning factor (Crews et al., 2001).

The content of 3-MCPD varies significantly among different edible oils. The content level, however, is strongly dependent on the formation potential, which is subject to the precursor content, as well as the technical parameters of the refining process. For palm oil, 3-MCPD-FE and related compounds can be detected at a level of up to 14 $mg \cdot kg^{-1}$ (FEI 2011, pp. 47–73; Zelinková et al., 2006).

On the contrary, natural and unrefined edible oils exhibit a significantly lower content of 3-MCPD due to the lack of intensive heat treatment during the production process.

A content of 74 $\mu g \cdot kg^{-1}$ of bound 3-MCPD was found in natural olive oil, while the refined product exhibited a content of 1464 $\mu g \cdot kg^{-1}$. Thus, the content in the refined product is approximately twenty times larger (19.78) compared to the natural product. Animal fats and oils are usually unrefined; therefore, it is not surprising that 3-MCPD and 3-MCPD-FE are not detectable in food products that are mainly based on animal fats and oils (Zelinková et al., 2006).

Finally, it should be noted that the content of 3-MCPD-FE exceeds the amount of free 3-MCPD in most of the tested samples by factor 396 (Svejkovska et al., 2004, pp. 190 - 196).

2.1.1.3 Appearance of 3-MCPD

Precursors of the formation of 3-MCPD in vegetable oils are lipids and chlorinated substances that are contained in the raw material, with the latter acting as chloride donators for the formation of 3-MCPD. This reaction was especially demonstrated during the refining process of palm oil. To initiate the reaction of lipids with chlorinated substances through heat, high temperatures are required.

Organic chlorinated substances are taken up by the palm tree via fertilizers and pesticides. During the plant growth and the production process of the crude oil, these compounds are transformed into mono chlorides. During the refining process of edible palm oil they react with lipids to 3-MCPD (Craft et al., 2011).

Also triacylglycerides (TAG) are important lipoid precursors which represent 90 to 95% of the the raw oil. Other lipids like glycerin, monoacylglycerides (MAG) and phospholipids are not considered as being essential for the heat induced formation of 3-MCPD during deodorization in practice. These compounds are eliminated before from the oil during the so-called 'de-gumming' process, which is an essential sub-process of plant oil refining. Diaclyglycerides (DAG) can also be active in the process of 3-MCPD-FE formation (Craft et al., 2013). With a content of at least 4 % in the raw oil DAG significantly contribute to the formation potential of 3-MCPD and related compounds like glycidol (Matthäus et al., 2011, pp. 380-386; FEI 2011, pp. 47–73).

For fruit pulp oils and palm oil, contents of DAG between 6 and 10% could be detected in the total raw material. This is a comparatively high level in comparison to other edible oils. The reason for this high level can be attributed to the activities of lipases (Bockisch, 1993, p. 463; Matthäus et al., 2011, pp. 380-386; Böttcher, 2007).

Beginning with a temperature between 120°C and 130°C, the organic chloride compounds of the raw oil are decomposed which results in free chloride ions. These intermediates are required for the formation process (Nagy et al., 2011) of 3-MCPD-FE, for which only four subsequent reactions need to be considered in a hydrophobic environment (Figure 2.2).

In all four cases the 3-MCPD-FE formation is initiated by a nucleophilic attack of a negatively charged chloride ion on the acylglycerin group in the TAG, DAG or MAG molecule.

In the first two reaction mechanisms (Figure 2.2 a/b), a direct nucleophilic attack of a chloride ion on a carbon atom in the glycerin backbone occurs for substances that carry either an ester group (a) or a hydroxyl group (b).

In the other two mechanisms (Figure 2.2 c/d) acylglycerin is produced as an intermediate in the first step of the process. In the second step of the process c, an acyloxonium ion is split from the acylglycerin. In the third step, the 3-MCP-FE molecule is generated while H_2O is split-off. In the alternative reaction d, an epoxide molecule is split from the acylglycerin in step 2 (Rahn and Yaylayan, 2011).
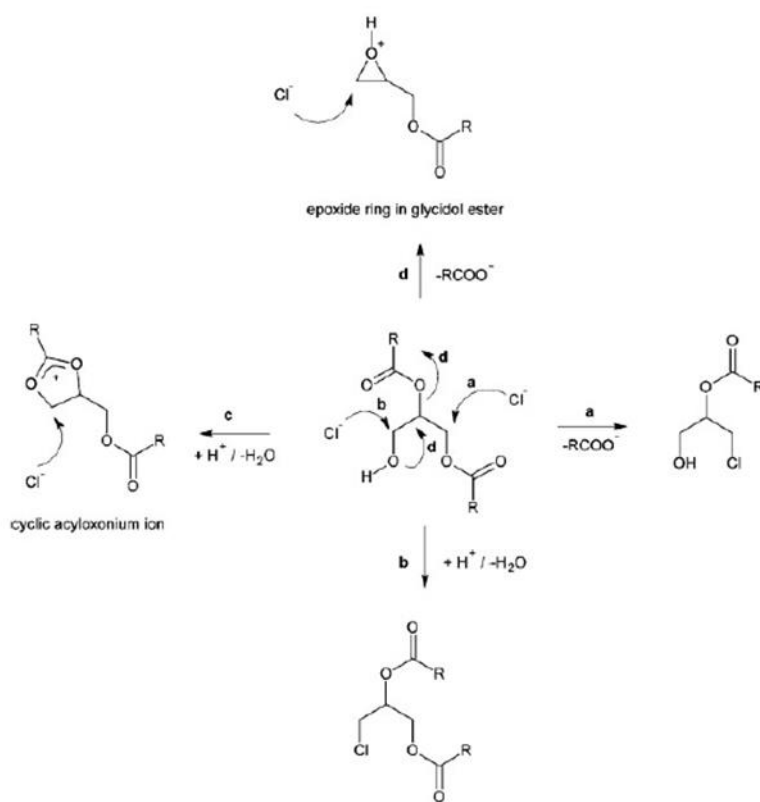
Figure 2.2: Alternative mechanisms for 3-MCPD-FE formation
a): direct nucleophilic attack of cloride ion on esterified glycerin carbon atom;
b): direct nucleophilic attack of chloride ion on hydroxylated glycerine carbon atom;
c): formation of cyclic acyloxonium ion from glycerin and subsequent nucleophilic attack of chlorine ion on carbon atom in oxonium ring;
d): formation of epoxide ring from glycerin and subsequent nucleophilic attack of chlorine ion on carbon atom in epoxide ring
(Rahn und Yaylayan, 2011)

The formation of 3-MCPD-FE in vegetable oils and vegetable fats is strongly dependent on environmental conditions. The thermal conditions of the refining process show a direct impact on the final 3-MCPD-FE content. In this regard the most important refining sub-process for the production of palm oil is the deodorization, where the highest temperatures of the entire refining process can be measured (Pudel et al., 2011; Zelinková et al., 2006).

There is a positive correlation between the temperature applied to the deodorization of palm oil and the overall 3-MCPD-FE content in the final product. For process temperatures below 200°C, the formation of 3-MCPD from triglycerides and chlorinated substances cannot be observed (Destaillats et al., 2012, pp. 29-37).

For the deodorization of bleached palm oil, an increase of the 3-MCPD-FE content from $4 \, \text{mg} \cdot \text{kg}^{-1}$ to $9 \, \text{mg} \cdot \text{kg}^{-1}$ can be observed when the process temperature is increased from 180°C to 250°C. Despite of this, the content of 3-MCPD-FE is reduced back to $6 \, \text{mg} \cdot \text{kg}^{-1}$ when the temperature is further increased to 280 °C. A potential reason for this decrease at elevated temperatures is thermal decomposition (Zulkurnain et al., 2012, pp. 799-805). Also an increase of the 3-MCPD-FE content could be observed for palm oil when the concentration of organic and inorganic chlorinated substances is increased (FEI 2011, pp. 47–73; Destaillats et al., 2012, pp. 29-37).

2.1.2 Glycidol

The glycidol molecule (2,3-epoxi-1-propanol) is formed in two mirror-inverted molecular forms. Glycidol is an organic chemical compound from the group of epoxides and alcohols. It is a flavourless, odourless and colourless liquid. In presence of water, glycidol is formed in a slow process (Synowietz, 1984).

2.1.2.1 Physico-chemical characterization of glycidol

The formation of 3-MCPD-FE occurs along with the formation of glycidol in its free form and Glycidyl-FE. Under the influence of heat, both reactions were observed in vegetable fats and oils, as well as in the foodstuffs produced from those materials.

Glycidol is an epoxide of glycerin. As in case of 3-MCPD, the free hydroxyl group of glycidol can become part of an ester bond with free fatty acids. Furthermore, 3-MCPD can be formed by reaction of glycidol with chlorine ions (see Figure 2.2).

2.1.2.2 Occurrence of glycidol

A detectable content of Glycidyl-FE and 3-MCPD-FE was observed in refined oils, with refined palm oil exhibiting the highest concentrations. Palm oil can reach a Glycidyl-FE concentration of up to $10 \, \text{mg} \cdot \text{kg}^{-1}$ after the refining process.

The presence of this substance could also be verified for foodstuffs containing refined oils, e.g. mayonnaise (Crews et al., 2013). The formation of Glycidyl-FE is initiated during the deodorization (sub-process of refining) when the temperature exceeds 200°C. Beyond that mark the concentration of Glycidyl-FE in palm oil increases proportionally with increasing temperature. E.g. for the temperature range of 240°C to 270°C, an above average increase of up to $45 \, mg \cdot kg^{-1}$ was demonstrated in experiments (FEI 2011, pp. 47–73).

The content of DAG in crude plant oil plays an essential role in the formation process of Glycidyl-FE. Crude oils with a high DAG content of 87% also showed a high Glycidyl-FE formation capacity in refining experiments. Moreover, the level of Glycidyl-FE was ten times higher than in refined oils produced from crude oils with a DAG content of 5%.

It is noteworthy that α-MAG can also act as a precursor in the Glycidyl-FE formation process. However, this substance plays only a subordinated role as its concentration in raw materials is usually negligible. Finally, TAG exhibit only a low potential for the formation of Glycidyl-FE (Destaillats et al., 2011; Craft et al., 2013).

2.1.2.3 Appearance of glycidol

The heat induced formation mechanism of Glycidyl-FE is shown in Figure 2.3 for a diacylglycerine molecule (DAG) and in Fig 2.4 for a monoacylglycerine molecule (α-MAG). The formation of Glycidyl-FE from DAG can be described as a heat induced intramolecular relocation; it is similar to the formation of 3-MCPD-FE.
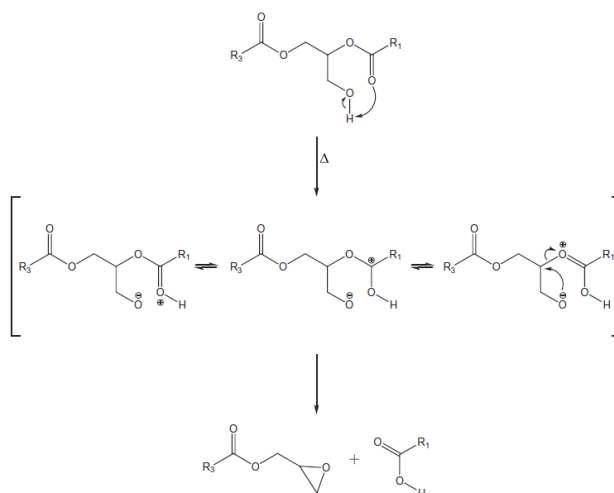
Figure 2.3: Glycidyl-FE formation starting with DAG molecule

The proton of the hydroxyl group is transferred to the neighboring oxygen atom that is contained in the esterified fatty acid. An acyloxonium ion occurs as reactive intermediate. By a nucleophilic attack of the negatively charged oxygen the oxiran ring is closed. Through the relocation of electrons, the fatty acid is split off the molecular backbone. The transformation of an α-MAG molecule to Glycidyl-FE is accompanied by a transfer of protons and a split-off of H_2O (Figure 2.4).
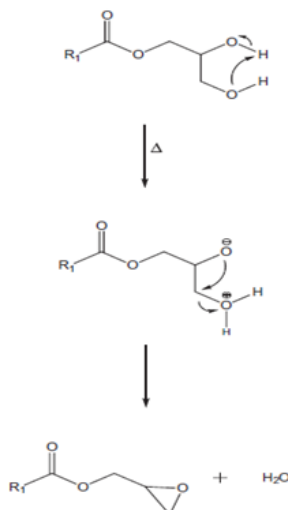
Figure 2.4: Glycidyl-FE formation starting with a MAG molecule

Finally, a 3-MCPD-monoester can act as a precursor in the Glycidyl-FE formation process. However, this reaction only plays a subordinate role in food production (Destaillats et al., 2011).

2.1.3 Relevance of 3-MCPD and glycidol for nutrition and food industry

Previous experiments using animals have proven the carcinogenic properties of free 3-MCPD and glycidol (Abraham et al., 2013; BfR, 2009). The presence of these substances in refined vegetable oils is therefore detrimental to the oil-producing and oil–processing industry. More than 90% of vegetable oils are dedicated to the production of foods and supplements for human nutrition. The crude oils are refined prior to further use in nutritional products in order to ensure the desired sensory and technological properties (Böttcher, 2007; OVID, YEAR). Free 3-MCPD is hereby classified as a 'potential human carcinogen' and the joint FAO/WHO Expert Committee on Food Additives defined its Tolerable Daily Intake (TDI) with $2\,\mu g \cdot (kg\,body\,weight)^{-1}$ (WHO et al., 2007).

Currently, no toxicological studies for bound 3-MCPD exist. But it is known, however, that bound 3-MCPD-FE is cleavaged in the human gostro-intestinal-tract into free 3-MCPD by 100%. Thus the same toxicity is assumed for the bound form as well (Abraham et al., 2013).

Although the majority of edible plant oil products is not heavily regulated, some products do have established limits for the amount of 3-MCPD that is allowed to be present. Two examples are soy sauce and hydrolysed vegetable protein (HVP) for which a critical value of $20\,\mu g \cdot kg^{-1}$ for 3-MCPD has been established by EU regulation no. 1881/2006 (EK, 2006).

Consequently, the threshold levels defined so far indicate, that the edible plant oil industry is only allowed to manufacture products that contain a maximum of $5\,mg \cdot kg^{-1}$ 3-MCPD. This regulation limit is based on an average adult with a bodyweight of 60 kg and a fat

consumption rate of $25\,\mathrm{g}\cdot\mathrm{per\,day}^{-1}$ (Gemeinschaftsausschuss für die Analytik von Fetten, 2012).

It is assumed that free glycidol also exhibits genotoxic and carcinogenic effects. As for 3-MCPD, it is supposed that Glycidyl-FE are completely cleavaged into free glycidol and fatty acids during the digestive processes.

As long as no threshold level or TDI is defined for glycidol or its fatty acid esters, the ALARA principle (As Low As Reasonably Achievable) is recommended. According to this principle the amount of Glycidyl-FE that is present in a product should be reduced to the smallest value possible. Through the use of this strategy the detrimental impact on humans is expected to be reduced to the lowest possible level (BfR, 2009).

Refined edible oils are also used as an essential ingredient in the production of baby milk powder and baby foods. The detection of health-critical doses of 3-MCPD in products intended for babies and infants negatively affects the image of the edible plant oil industry. A study on twenty food products found doses ranging from $62\,\mathrm{\mu g\cdot kg}^{-1}$ to $588\,\mathrm{\mu g\cdot kg}^{-1}$ 3-MCPD (Crews et al., 2002; Zelinková et al., 2009). According to a model calculation, these doses would exceed the TDI for a baby by a factor of twenty (BfR, 2007) under the following two conditions: First, the baby only consumes nutrients from foods designed for infants. Second, the 3-MCPD is released from the 3-MCPD-FE during production. In this study, it was also assumed that the TDI defined for adults is also valid for babies without a correction factor.

2.1.4 Minimizing strategies for 3-MCPD-FE und glycidyl-FE in plant oils

This subchapter gives an overview on ten technological strategies that can be used to reduce the content of 3-MCPD-FE and Glycidyl-FE in refined edible plant oils.

The first approach is to replace crude oil types with high precursor concentrations (e.g. palm oil) by those with lower ones (e.g. rapeseed oil). This approach is already in practice in the refining industry. However, it is not a feasible approach for a broad range of products. Especially palm oil exhibits major technological and economic adavantages which make it superior to other oils with respect to food production. Furthermore, the crop yield for palm tree plantations is $4{,}2\,\mathrm{t\cdot ha}^{-1}$, which greatly exceeds the crop yields for sunflower, rapeseed and soy ($0{,}69\,\mathrm{t\cdot ha}^{-1}$ on average). Also the demand for edible oils will rise by 500% until 2050 according to recent studies. Therefore, palm oil is of particular interest for feeding the world's population in the future (Junker, 2011; BVE, 2010). These aspects show that it is of special importance to identify contaminant reduction strategies for oil types with high precursor contents and formation potential, rather than just replacing one oil type by another.

A second approach is to improve the quality of crude oil, which has a deep impact on the content of toxic substances in refined oil down the production line. Quality parameters in this sense are the content of free fatty acids (FFA), phosphor, DAG and MAG in the crude oil. When low quality crude oil is used in the refining process, the level of 3-MCPD may be up to five times higher in the end product, compared to high quality crude oil. The level of Glycidyl-FE in the refined product is also positively correlated with the DAG content in the crude oil. Thus, manufacturers typically prefer crude oil of superior quality.

The DAG and FFA contents of the crude oil are also dependent on harvest and storage conditions, as well as the production method. All three factors strongly influence the formation of contaminants. Furthermore, late harvesting and extended storage times, as well as damage to the oil fruits, may cause enzymatic reactions which consequently cause lipid split-offs. Therefore, processing immediately after harvest, and thus limiting storage time, is important for the quality of the refined final product (Craft et al., 2012; Zulkurnain et al., 2012, pp. 799-805).

The third approach is to reduce the content of inorganic chlorinated compounds. Organic chlorinated compounds are synthesized by the plant metabolism from inorganic chlorinated compounds that have been absorbed before (e.g. chlorinated pesticides and fertilizers). These organic chlorinated compounds act as a chloride donor for the subsequent 3-MCPD formation. Therefore, the reduction of the content of inorganic chlorinated compounds is an effective strategy for improving the quality of the refined final product. Avoiding chlorinated pesticides and fertilizers during cultivation thus improves the overall quality of the end product. Furthermore, improving water quality is also helpful in this regard (Craft et al., 2011; FEI, 2011, pp. 47–73).

The fourth approach is to wash out chlorinated compounds by a washing step prior to the refining process. De-ionized water, chlorine free alcohol and hexan can be used for removal of these precursors. Through this method, the 3-MCPD content in refined plant oils can be reduced by one-third (Craft et al., 2011; FEI, 2011, pp. 47–73).

The fifth alternative is the application of two modified refining sub-processes, de-gumming (the removal of gums from the oil) and bleaching, prior to deodorization. Both process steps can lead to a reduction in the content of 3-MCPD-FE and related compounds by 40%. If acids are not used after the de-gumming step, the 3-MCPD content can be further reduced; in refined palm oil this reduction effect may range from a maximum of $2\,\mathrm{mg\cdot kg^{-1}}$ down to about $0,1\,\mathrm{mg\cdot kg^{-1}}$ (factor twenty).

Alternative six is to apply acid-activated bleaching earth to the refining process. It is assumed that the acid-based protonation of lipids inhibits the formation of reactive precursors. An alteration of the refining conditions in this context can only be hardly introduced by the

industry. The reason is that by this approach the oil quality is suffering. Oil quality is, however, another important factor in the whole production process (FEI, 2011, pp. 47–73; Ramli et al., 2011; Zulkurnain et al., 2012, pp. 799-805).

Alternative number seven is to add auxiliary materials to the deodorization and to treat the refined product with absorbers in an effort to reduce 3-MCPD-FE and related substances to a minimum. Although this approach is considered to be effective, it is still under development and actually not ready for industrial-scale implementation (FEI, 2011, pp. 47–73).

The eigth alternative takes into consideration that the generation of 3-MCPD-FE und Glycidyl-FE is induced by heat during the production process. The formation of these compounds only takes place in the upper heat region of the deodorization process. Thus, the contaminant formation in the final product can be controlled by reducing the maximum temperatures that are applied to the deodorization process. Unfortunatly this approach conflicts with the goal of fulfilling other oil quality criteria such as acid value or oxidation stability (Pudel et al., 2011; FEI, 2011, pp. 47–73).

The ninth alternative is to apply a two-step deodorization (two successive standard deodorization processes with different duration and temperature). In comparison to a one-step method, it can reduce the content of 3-MCPD-FE by up to 80%. With this approach, all quality criteria are fulfilled. The process for industrial use of the two-step deodorization is under development.

The present thesis focuses on the tenth approach, which is the Short-Path-Distillation (SPD), and its potential advantages for the refining industry. The SPD is a sophisticated and technically highly complex process that enables the reduction of contaminant formation, along with satisfactory oil quality, by operating at fine vacuum and with special technical characteristics (FEI, 2011, pp. 47–73).

2.2 Principle of Short-Path-Distillation (SPD)

Standard deodorization in plant oil refining has been identified as being crucial for the heat induced formation of 3-MCPD and Glycidyl-FE due to a heat exposition of up to 120 min at 260°C - 270°C (Pudel et al., 2011). Recent research results (FEI, 2011) indicate the potential of the SPD to replace the standard deodorization regarding the control of contaminant formation while ensuring oil quality. A process has been already developed by which bleached palm oil has been deodorized successfully at a temperature of 170°C and a pressure of 0,13 mbar concomitant with meeting the standard quality specifications. The formation of 3-MCPD- and G-FE was not considered here (Ooi et al., 1996).

In this thesis a detailed investigation and process optimization of palm oil Short-Path-Distillation (SPD) towards contaminant formation and standard industrial quality parameters is presented.

Specific technological characteristics of the SPD enable the efficient distillative separation of volatile compounds from a liquid multicomponent mixture at comparatively mild conditions:

- Vacuum down to 10^{-3} mbar (lowers the boiling point)
- Low residence time of oil volume unit on evaporator surface (reduced heat exposition)
- Evaporation from a thin liquid layer
- Turbulent recirculation of the film by agitation
- Short path between evaporator and condenser [1]

Especially the short path optimizes separation by avoiding re-condensation at the evaporator surface due to collision of molecules. Here the laboratory plant VKL-70-5 manufactured by the company VTA GmbH has been applied (VTA, 2010). This plant is characterized by a double-wall glas cylinder as evaporator with an inside axially arranged glas spiral as condenser. Both can be adjustably tempered from inside via thermo oil by a thermostat. The pre-tempered oil is fed from the tank to the inner evaporator surface from the top of the evaporator by an adjustable pump. Radially arranged rolls dispense the oil on the inner evaporator surface as thin film whose thickness and turbulent recirculation can be influenced by the adjustable rotator speed. The distillate (substances to be separated) flows down the condenser and the residue (refined oil) down the evaporator surface into separate tempered receiver tanks (see also subchapter 4.1).

2.3 Model based process optimization

For most technological processes the influence of process variables (e.g. temperature) and initial conditions (e.g. initial educt concentration) on the process output (e.g. quality criteria of a product) is of high technological and physico-chemical complexity. Thus, unsystimatical and arbitrary (e.g. trial-and-error) process adjustment approaches are not suitable for a time and cost efficient identification of process settings that meat the optimization criteria of the process output. Mathematical process models simulate the effects of the most significant process variables and initial conditions on the target values (product quality parameters) over the whole range of reasonable adjustments. The model terms and respective constants are determined systematically by a limited number of specific optimized experimental settings (control of experimental effort). This enables the cost-efficient prediction of process

---

[1] It is below the mean free path of the evaporated compounds at this pressure.

outcomes at process settings which have not been experimentally tested and, thus, the identification of process settings that fulfill the optimization goal.

2.3.1 Comparison of mechanistic and stochastic approaches

Mechanistic models aim to picture the dynamics and kinetics of distinct physico-chemical and technological sub-processes by appropriate mathematical expressions. The required characteristic apparative, kinetic and material constants are determined experimentally or taken from the literature. These models can be easily adapted to different process setups by just adapting the values of the respective constants. Therefore an appropriate mechanistic process modeling requires detailed knowledge on all relevant sub-processes. Statistical modelling is the number one choice for the design and control of processes in many industries today. The primary advantage towards mechanistic modeling is that no deeper knowledge on technological and physico-chemical sub-processes is required and also in many cases a more precise simulation is provided as all potential effects and interactions can be theoretically considered. Unfortunately, statistic models are only valid for the process setup they were determined at and there is no technological and physico-chemical means of defining the parameters in such models. Mechanistic models, however, are usually based on the laws of physics, chemistry and process engineering. Therefore, the parameters of such models are able to be defined in a much more concrete manner and adaptable to alterations in process setups (Staby, Arne, 2014).

2.3.2 Stochastic approaches

The measured experimental data is affected by a randomly occurring error in experimentation which is in most cases Gaussian distributed. Thus it is unknown to what extent the measured data set, and therefore also the model and its parameterization calculated from that data, approach the real underlying process.

Given a defined underlying process and error of experimentation, a broad range of different experimental data sets can occur with different probability. As model selection, model parameterization and model solutions are a function of the experimental data, they approach the real underlying process only with a certain probability. The higher the error of experimentation is, the larger is the range of possible results and the lower is the reliability of results. That is true for model parameterization and solutions that can occur for a defined underlying process. This makes us aware of the fact that process modelling is always accompanied by uncertainties in model selection, parameterization and solution. Consequently a comprehensive stochastic evaluation of model performance is required, which evaluates these factors and their validity in terms of probability and range of possible alternatives.

2.3.2.1 Statistic Design of Experiments and Response Surface Methodology

This chapter describes the principles of the Response Surface Methodology (RSM). RSM is based on Statistical Design of Experiments (SDoE) and is an approach which includes experimental designs of a test plan of different complexity which in turn enable the derivation of linear process models of different complexity. The less complex full factorial design enables the derivation of process models of 1st grade. Specifically, 1st grade means that the model is based on a degenerated polynomial that comprises all linear effects of the process variables towards the target value as well as all possible interactions (up to highest grade). This is followed by 2nd grade models which are determined from more complex Central Composite Designs (CCD). Such models are based on full polynomials of 2nd grade that comprise all linear effects and quadratic effects of the process variables towards the target value as well as interactions up to 1st grade (two-factor-interactions). Both experimental designs, models respectively, will be applied and derived in this thesis for the model-based optimization of the Short-Path-Distillation (SPD).

2.3.2.1.1 Principles and targets of RSM

In order to optimize production processes, informations on the influence of process variables on the process output are required. In this regard, evaporator temperature and rotation speed of the stirrer are examples for process variables that are relevant for the SPD. Product quality parameter (process output) for an SPD process applied to palm oil deodorization could be the content of tocopherol or the acid value, for example. Moreover, the acid value is the amount of a base $(mg \cdot (g\ oil)^{-1})$ that is required to neutralize any acids that are contained in the oil.

The effect of the process variables on the target values is analyzed and optimized model based. As the deodorization of bleached palm oil by SPD is characterized by very complex physico-chemical and technological processes, mechanistic modeling approaches are limited by complexity and insufficient knowledge. The RSM is a self-contained stochastic approach that models the effects and interactions of process variables towards a target value without considering the physico-chemical and technological sub-processes being responsible for that effects. This so called "black box" approach models the output (target values) just as function of the input (process variable) by corresponding terms in a linear model that sums up the effects and interactions of process variables towards a defined target value. This derivation of a process model is based on a Statistical Design of Experiments (SDoE). By this approach the minimum number of experimental settings that is necessary for model derivation is defined by mathematical means. This optimization of experimental settings represents the minimization of the experimental effort necessary to calculate the single effects and interactions to be considered in the model. Thus, the information on the true underlying

process contained in the experimental data is optimized. Generally, models of different complexity can be derived based on experimental designs of different complexity. Here models of 1st and 2nd grade are derived and compared in performance. For each target value the dependency on the process variables to be considered is modeled. The necessary adjustments of the process variables to meet product quality requirements can be predicted by such mathematical models. Thus, RSM enables the optimization of one or more process variables simultaneously based on process models.

For the deodorization of vegetable oils by SPD, it can be observed that process variables influence the process output in multiple ways. The derivation of mathematical models enables the simulation of these effects. In this context, RSM is a cost efficient approach which also provides graphical presentation of the process dynamics as basis for optimization approaches (Box and Draper, 2007, p. 1). Therefore the combination of SPD and RSM is a promising approach for optimizing the deodorization process.

2.3.2.1.2 Principles and targets of SDoE

Initial statistical pilot studies are applied in SDoE to determine the overall feasibility of a potential experiment with respect to a certain goal (e.g. identification of significant effects) and, thus, are the basis for SDoE. Variable factors of experimental settings, along with the exact number of experiments, are fixed ex ante here. Each experiment must be executed as part of a whole set of experiments, which is the test plan of the SDoE. Moreover, each target value is determined for each trial separately.

Multiple statistically meaningful ways exist for the SDoE, all of which vary in the number of experiments required and the complexity of the mathematical models derived from these experiments.

The basic rules for modeling are, that for more complex models also more complex designs with a higher number of experiments are required and that the prediction quality increases with the number of process variable effects and interactions that are considered in the model.

It should be kept in mind, however, that the additional effort (cost) for the experiments is not always justified by an increase in accuracy. When the target of a design is to optimize multiple goals, a separate model can be desiged for each individual goal. This approach is known as 'analysis and planning under multiple objectives' (Scheffler, 1986, pp. 11–46).

When experiments are complex and cost intensive (e.g., in the case of SPD) SDoE is the method of choice. With the minimum number of optimized trials required, a maximum of information on process dynamics can be gained and the whole process can be optimized through the incorporation of models derived from the SDoE. Specifically, product quality is improved and costs are reduced (Kleppmann, 2008, pp. 1–24; Scheffler, 1986, pp. 11–46).

2.3.2.1.3 2^m full factorial design and models of 1st grade

Each process variable, or factor, in a full factorial design for the derivation of a 1st degree model exhibits exactly two settings, x_{max} and x_{min}, which are the maximum and minimum values that can be reasonably achieved within the range considered. Each process variable range can be defined individually. The following questions aid this definition: 'What happens in practice?' and 'What are the limiting factors of my test facility?'

Test plans for mathematical models of higher degrees exhibit more than two values to be achieved for each factor. In a full factorial design of type 2^m with m process variables, the test plan includes exactly $n = 2^m$ experiments. All possible combinations of process variable settings must be tested (see Table 2.1). The experimental results indicate the influence of each analyzed process variable towards the target values measured. This in turn enables the derivation of linear models and a model based optimization of the process in the next step (Scheffler, 1986, pp. 11–46).

A 2^m full factorial design enables the researcher to calculate the individual linear effects of m factors (process variables) towards the target values measured for each experiment. Interactions between factors up to highest grade can also be calculated. The mathematical model is derived based on a degenerated polynomial that comprises these linear effects and interactions. Eq. 2.1 exhibits the general form, Eq. 2.2 a model for four process variables as applied here for the SPD.

The limitation to 2^m full factorial designs and 1st grade models is not feasible for more complex problems. To identify and model also effects of quadratic or higher order more than two values for each factor must be analyzed in a more complex SDoE, thus enabling the derivation of more complex models.

$$\hat{y} = \beta_0 + \sum_{j=1}^{m} \beta_j \cdot x_j + \sum_{j=1}^{m}\sum_{k=j+1}^{m} \beta_{jk} \cdot x_j \cdot x_k + \sum_{j}^{m}\sum_{k=j+1}^{m}\sum_{l=k+1}^{m} \beta_{jkl} \cdot x_j \cdot x_k \cdot x_l \ \dots.$$

Equation 2.1: Polynomial for model of 1st grade (general form)

From the test plan and the referring experimental results also the interdependencies between factors (exposed in Eq. 2.1 for the general case and in Eq. 2.2 for 4 factors) can be calculated. For all possible combinations of input values x_j, target value $\hat{y}$ can be calculated by Eq. 2.1 and Eq. 2.2 (NIST/SEMATECH, 2012; Scheffler, 1986, pp. 32–36; Kleppmann, 2008, pp. 24, 198).

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4$$
$$+ \beta_{34} x_3 x_4 + \beta_{123} x_1 x_2 x_3 + \beta_{124} x_1 x_2 x_4 + \beta_{234} x_2 x_3 x_4 + \beta_{1234} x_1 x_2 x_3 x_4$$

Equation 2.2: Model of 1st grade for 4 factors

$\hat{y}$ is the target value calculated by the model and x_j is the adjustment of process variable j in normalized form. The one-dimensional coefficients β_j are estimations for the model parameters of the corresponding linear effects. These coefficients are determined from the

experimental results referring to the experiments of the test plan. They quantify the change in the target value that is caused by the change of the normalized process variable from a medium adjustment to ±1.

In the same manner, the coefficients β_{jk} and β_{jkl}, for example, contain more than one dimension as they quantify the interactions between two and three process variables regarding their effect on the target value. The constant β_0 is the effect that takes place when all process variables of the model are adjusted to the mean value of the analysis range. A correlation (COR) or interaction between process variables (factors respectively) is present when the effect of one factor towards the target value is dependent on the adjustment of one or more other factors. In this context, e.g. the term $\beta_{12}x_1x_2$ in Eq. 2.2 represents a two factor correlation (2FCOR). With COR, it is possible to gain information from a model about the reciprocal influences of process variables. This enables the interpreter to gain insight into physico-chemical processes and further aids the development and scale-up of production processes (Scheffler, 1986, pp. 15-19; 32-36; Kleppmann, 2008, pp. 24, 99, 198).

Figure 2.4 graphically illustrates a full factorial experimental design of type 2^3. The three dimensions / axes of the design represent the three factors (x_1, x_2, x_3) with two normalized values each (max = +1 and min = -1).
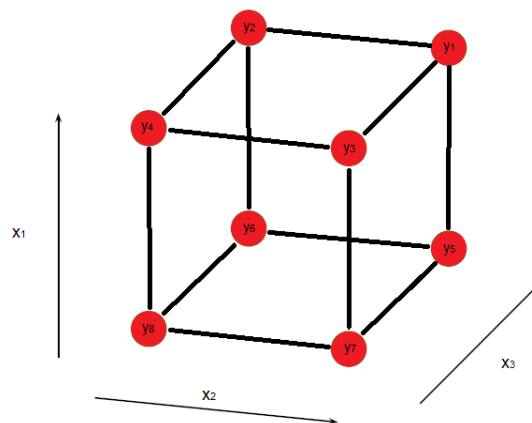
Figure 2.5: 2^3 full factorial design (Siebertz et al. 2010, p. 8)

Table 2.1 shows all normalized values for the process variables (factors) of the process for a 2^3 full factorial design. The variable y relates to the experimental results for the target values to be optimized which are measured for the experimental settings represented by each corner point of the cube in Figure 2.5 (Siebertz et al. 2010, p. 7).

Experiment	x_1	x_2	x_3	Result
1	+1	+1	+1	y_1
2	-1	+1	+1	y_2
3	+1	-1	+1	y_3
4	-1	-1	+1	y_4
5	+1	+1	-1	y_5
6	-1	+1	-1	y_6
7	+1	-1	-1	y_7
8	-1	-1	-1	y_8

Table 2.1: 2^3 full factorial design with three factors on two normalized levels each

A model of 1st grade for four process variables, as applied for analysis of SPD, considers single effects and interactions up to 3rd grade (see Eq. 2.2). It is based upon a 2^4 full factorial design (see Table 2.2) with n=16 experiments.

experiment	X_0	X_1	X_2	X_3	X_4	$X_1 \cdot X_2$	$X_1 \cdot X_3$	$X_1 \cdot X_4$	$X_2 \cdot X_3$	$X_2 \cdot X_4$	$X_3 \cdot X_4$	$X_1 \cdot X_2 \cdot X_3$	$X_1 \cdot X_2 \cdot X_4$	$X_1 \cdot X_3 \cdot X_4$	$X_2 \cdot X_3 \cdot X_4$	$X_1 \cdot X_2 \cdot X_3 \cdot X_4$
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	-1	-1
3	1	1	1	-1	1	1	-1	1	-1	1	-1	-1	1	-1	-1	-1
4	1	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1
5	1	1	-1	1	1	-1	1	1	-1	-1	1	-1	-1	1	-1	-1
6	1	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1
7	1	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1
8	1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1
9	1	-1	1	1	1	-1	-1	-1	1	1	1	-1	-1	-1	1	-1
10	1	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1
11	1	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1
12	1	-1	1	-1	-1	-1	1	1	-1	-1	1	1	1	-1	1	-1
13	1	-1	-1	1	1	1	-1	-1	-1	-1	1	1	1	-1	-1	1
14	1	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1
15	1	-1	-1	-1	1	1	1	-1	1	-1	-1	-1	1	1	1	-1
16	1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1

Table 2.2: Trial and calculation matrix for a 2^4 full factorial design

The trial matrix in Table 2.2 contains the settings of the process variables for the experiments, where 1 and -1 are the normalized minimum and maximum of the variable range analyzed. This table also shows the normalized settings of the interactions which are the product of the referring single variable settings. These normalized interaction settings, as well as the 'setting' of the constant model term in the second column of Tabel 2.2, are not of practical relevance for experimentation but only for the calculation of model parameter values from the experimental results.

Normalization is achieved according to Eq. 2.3 where $\widetilde{x}_{ij}$ is the normalization of the ith setting of process variable j and x_{ij} is the non-normalized setting of variable j (Eq. 2.4).

$$\widetilde{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{x_{max,j} - \bar{x}_j}$$

Equation 2.3: Normalization of factor values

$\bar{x}_j$ is the arithmetic mean of the value $x_{min,j}$ and $x_{max,j}$ (Eq. 2.4).

$$\bar{x}_j = \frac{x_{max,j} + x_{min,j}}{2}$$

Equation 2.4: Calculation of arithmetic mean

From the results of all n experiments, the parameters β_j of the 1st grade model (Eq. 2.1 and Eq. 2.2) are calculated according to Eq. 2.5, where y_i is the experimental result for a specific target value of the ith experiment and $f_j(x_i)$ is the normalized setting of the jth process variable (factor respectively) in the ith experiment (Bandemer et al., 1973, p. 9).

$$\beta_j = \frac{\sum_{i=1}^{n} f_j(x_i)y_i}{n}$$

Equation 2.5: Estimation of parameters β_j

Significant parameters / model terms are identified by statistical significance analysis. For this purpose it is assumed that the interactions of higher than 2nd order are quantitatively or theoretically reasonable negligible and thus represent an estimate for the error variance of experimentation. The validity of these interactions as such an estimate is statistically proven by application of the Bartlet-criterion. The significance of the remaining parameters is then determined by calculation of the ratio between variance of the referring effect (or interaction) and the error variance, followed by statistical assessment of significance via F-test. Thus, statistical significant effects can be identified and summarized to a potential model candidate. By that procedure a model candidate of 1st grade can be derived for each target value considered. The model adequacy is tested for each model candidate by calculation of the relative variance according to Eq. 2.6 followed by a statistical assessment via F-test. In Eq. 2.6 n is the number of experiments of the SDoE, p is the number of model parameters and n_T is the total number of experiments achieved (differs from n in case of multiple repetition of SDoE). The summed squares QS_L and QS_F are calculated according Eq. 2.7 and 2.8 where n_i is the number of repetitions of the ith experiment of the test plan, $\bar{y}_i$ is the average result of the ith experiment, $\hat{y}_i$ is the model solution at the process variable setting of the ith experiment and y_{il} is the result of the lth repetition of the ith experiment.

$$v = \frac{\dfrac{QS_L}{n-p}}{\dfrac{QS_F}{n_T-n}}$$

Equation 2.6: Relative variance for model adequacy testing

$$QS_L = \sum_{i=1}^{n} n_i(\bar{y}_i - \hat{y}_i)^2$$

Equation 2.7: Summed squares QS_L

$$QS_F = \sum_{i=1}^{n} \sum_{l=1}^{n} (y_{il} - \bar{y}_i)^2$$

Equation 2.8: Summed Squares QS_F

2.3.2.1.4 Central composite design (CCD) and models of 2nd grade

A statistical linear model of 1st grade comprises linear effects of the process variables towards the target value, as well as interdependencies of 1st grade and higher order. In many technical processes, the effects of process variables are the result of more complex underlying physico-chemical sub-processes. Furthermore, their curve progression most likely contains exponential characteristics. If the proportion of exponential parts in the curve is limited, linear models can typically approach a process with sufficient accuracy. However, if a process exhibits significant non-linear characteristics in this regard, a model considering square effects becomes the method of choice. A polynomial of 2nd degree is suitable to describe the square effects of process variables towards the target values. With these models, local maxima and minima for the target value can be calculated.

$$\hat{y} = \beta_0 + \sum_{j=1}^{m} \beta_j \cdot x_j + \sum_{j=1}^{m} \beta_{jj} \cdot x_j^2 + \sum_{j=1}^{m} \sum_{k=j+1}^{m} \beta_{jk} \cdot x_j \cdot x_k$$

Equation 2.9: Polynomial of 2nd grade (general form)

Eq. 2.10 demonstrates the application of the general 2nd grade polynomial (see Eq. 2.9) to an example with four factors.

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{44} x_4^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3$$
$$+ \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4$$

Equation 2.10: Polynomial of 2nd grade for four factors

The more complex a model is, the more experimental data input is needed to estimate its parameters. Thus, the system dynamics of a 2nd grade model can be described by the results of additional experimental trials compared to the 1st grade model (full factorial design respectively). For the resolution of all linear and quadratic effects, as well as all interactions up to highest order, a 3^m full factorial design is required. In such design, factors must be adjusted to three settings. However, this subset of test plans usually demands an enormous test effort that cannot be commercially justified. For this reason they are rarely used in practice.

Another subset of plans is the Central Composite Design (CCD), which enables the determination of the parameters in a reduced 2nd grade model, which is based on a polynomial of 2nd grade (Eq. 2.9 and Eq. 2.10), with less experimental effort compared to 3^m full factorial designs (Klein, 2011, p. 150 and Scheffler, 1986, p. 201–208). CCDs represent

an enhancement of 2^m full factorial designs by specific mathematically defined additional experiments which enable the resolution of the reduced model of 2nd grade exposed in Eq. 2.9 and 2.10. These designs include a 2^m full factor design in their core which is enhanced by the star points and the central point of the core (at this point, all normalized process variables are adjusted to zero). Thus the central point can be also interpreted as the average point of all factors (Kleppmann, 2008, p. 198). This enhancement of the core mathematically enables the identification and estimation of squared effects of the process variables on the target values while interactions of higher than 1st order are neglected (compare Eq. 2.9 and 2.10). Figure 2.6 graphically exposes a CCD for three factors.

The total number of experiments of a CCD is denoted as N_{tot} and defined in Eq. 2.11.This value sums up the number of trials for the core (N_w), the number of trials for the starpoints (N_{St}) and the number of trials for the center point (N_0).

$$N_{tot} = N_W + N_{St} + N_0 = 2^m + 2m + 1$$

Equation 2.11: Number of trials in CCD

The derivation of Eq. 2.5 for the estimation of parameters of 1st grade models is based on symmetry (all experimental settings are symmetric with respect to the center) and orthogonality of the calculation matrix as basic requirement. To enable an analogous estimation of parameters of a 2nd grade model (compare Eq. 2.9 and 2.10) from the corresponding calculation matrix of a CCD, also this matrix has to exhibit orthogonal properties which are provided by coordinate transformation. The latter results in the orthogonal calculation matrix (orthogonal CCD respectively) exposed in Table 2.3 and the corresponding transformed 2nd grade model given by Eq. 2.15 (for the general case). Eq. 2.14 exposes the calculation of the transformation factor and Eq. 2.13 the calculation of the extension factor where q is the index of partial designs which is zero in case of full central composite designs (as applied in this work for the analysis of the SPD).

The total number of trials needed for an orthogonal CCD is dependent on the parameters c_w, c_{St} and c_o of the variables N_0, N_{St} and N_W in Eq. 2.12 (Scheffler, 1986, p. 224).

$$n_{tot} = n_W + n_{St} + n_0 = c_W + c_{St}N_{St} + c_0N_0$$

Equation 2.12: Total number of trials in orthogonal CCD

For the analysis of quadratic effects by an orthogonal CCD, each factor is tested on three additional levels. The total number of levels for each factor is therefore increased to five (± 1, $\pm \delta$ and 0; compare Figure 2.6). In orthogonal CCDs, star points have a distance of δ to the center point (Siebertz et al., 2010, p. 38–40 and NIST/SEMATECH, 2012).

$$\delta^2 = \sqrt{2^{\frac{m-q}{2}-1}\left(\sqrt{n-2^{\frac{m-q}{2}}}\right)}$$

Equation 2.13: Calculation of extension factor

$$\gamma = \frac{2^{m-q} + 2\delta^2}{n}$$

Equation 2.14: Calculation of coordinate transformation factor
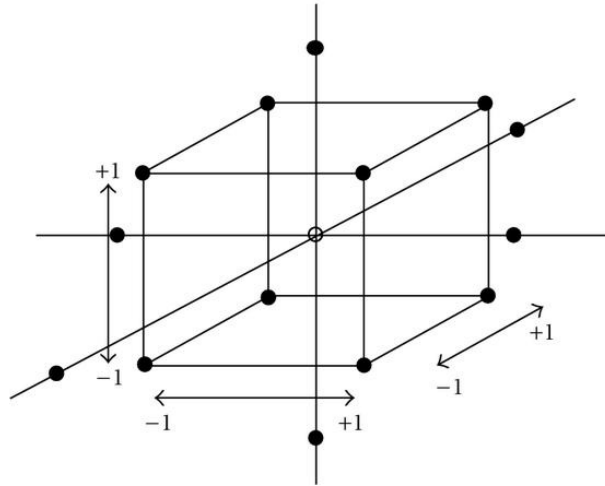
Figure 2.6: Orthogonal CCD for 3 factors with central and star points (Siebertz et al. 2010)

Figure 2.6 graphically exposes an orthogonal CCD for 3 process variables. Table 2.3 exhibits the calculation and trial matrix of an orthogonal CCD for the general case. The parameter values of the non-transformed model (Eq. 2.9 and Eq. 2.10) are calculated from the experimental results and the calculation matrix according to Eq. 2.16 and 2.17 where m is the number of process variables, x_{ij} the normalized setting of variable j in the ith experiment and y_i the measured result for the target value from the ith experiment. The values for c, δ and γ are calculated by complex mathematical operations according to Eq. 2.13, 2.14 and 2.17. They can be also taken from tables in the literature for common cases (Scheffler, 1986, p. 222-244). Parameter significances are again determined by calculation of the ratio between the variance of the referring effect (or interaction) and the error variance. The error variance can be estimate from negligible model terms (as exposed for models of 1st degree), the initial assumption that the tested model is adequate (the validity of this assumption is proofed subsequently) or from multiple repetitions of experiments. For each of these alternatives the significance of model terms is tested by a F-test of the relative variance (as exposed for models of 1st degree) or t-test of the individual confidence intervals. Thus six different ways for the determination of significant model terms are possible and achieved here which might produce different models of 2nd degree for a

certain target value. These statistical methods are considered as basic knowledge of statistics and are therefore not further exposed in detail here.

$$y = a_1 + \sum_{j=1}^{m} \beta_j x_j + \sum_{j=1}^{m} \beta_{jj}(x_j^2 - \gamma) + \sum_{j=1}^{m} \sum_{k=j+1}^{n} \beta_{jk} x_j x_k$$

Equation 2.15: Transformed model of 2nd grade (general form)

$$ß_0 = a_1 - \gamma \cdot \sum_{j=1}^{m} b_{jj} \ with \ a_1 = c_1 \cdot \sum_{i=1}^{n} Y_i$$

$$ß_j = c_2 \cdot \sum_{i=1}^{n} x_{ij} \cdot Y_i \ for \ j = 2, \dots, m+1$$

$$ß_j = c_3 \cdot \sum_{i=1}^{n} (x_{ij}^2 - \gamma) \cdot Y_i \ for \ j = m+2, \dots, 2m+1$$

$$ß_j = c_4 \cdot \sum_{i=1}^{n} x_{ij} \cdot x_{ik} \cdot Y_i \ for \ j = 2m+2, \dots, p$$

Equation 2.16: Alternative formulas for the calculation of model parameters $ß_j$

$$c_j = m_j^{-1}$$

$$with \ m_1 = n, m_2 = n\gamma, m_3 = 2^{m-q}(1-\gamma)^2 + 2(\delta^2 - \gamma)^2 + (2m-1)\gamma^2, m_4 = 2^{m-q}$$

Equation 2.17: Values c_j for the calculation of model parameters

Trial number	Calculation matrix											Subset
	Trial matrix (experiments)											
	X_0	X_1	X_2	...	X_m	$X_1^2-\gamma$	...	$X_m^2-\gamma$	X_1X_2	...	$X_{m-1}X_m$	
1	1	-1	-1	...	-1	$1-\gamma$	...	$1-\gamma$	1		1	Core: 2^{m-q} factor plan
2	1	1	-1		-1	$1-\gamma$	...	$1-\gamma$	-1		1	
⋮	⋮					⋮		⋮				
2^{m-q}	1					$1-\gamma$	...	$1-\gamma$				
$2^{m-q}+1$	1	δ	0	...	0	$\delta^2-\gamma$		$-\gamma$	0	...	0	Star points
$2^{m-q}+2$	1	-δ	0	...	0	$\delta^2-\gamma$		$-\gamma$	0	...	0	
	1	0	δ		0	$-\gamma$		$-\gamma$	0	...	0	
	⋮	⋮	-δ		0	$-\gamma$		$-\gamma$	0	...	0	
			⋮		⋮	⋮			⋮		⋮	
	1	0	0	...	0	$-\gamma$		$-\gamma$	0	...	0	
	1	0	0	...	δ	$-\gamma$		$\delta^2-\gamma$	0	...	0	
$2^{m-q}+2m$	1	0	0	...	-δ	$-\gamma$		$\delta^2-\gamma$	0	...	0	
$n=2^{m-q}+2m+1$	1	0	0	...	0	$-\gamma$		$-\gamma$	0	...	0	Center point

Table 2.3: Orthogonal CCD (general case)

2.3.2.2 Bayesian model fit, discrimination and validation

By the procedure exposed so far, model candidates of 1st and 2nd grade are derived by RSM for each analyzed target value of the SPD. Each model comprises the effects and interactions of the process variables towards a specific target value, which have been identified by RSM as being significant for the process. These alternative models are further analyzed and assassed by sophisticated Bayesian approaches which comprise model fit, discrimination, validation and performance analysis.

2.3.2.2.1 Bayesian concept

The Bayesian theory on model parameterization (fit respectively), discrimination, evaluation and validation provides the required measures of result uncertainty by means of probability distributions over the range of possible results. The Bayesian law of conditional probabilities, which is the basis of the Bayesian methods applied here, is a strictly probabilistic approach that provides the mathematical linkage between inverse conditional probabilities (Eq. 2.18).

$$p(\theta|y, M) = \frac{p(y|\theta, M) \cdot p(\theta|M)}{p(y|M)}$$

Equation 2.18: Bayesian law of conditional probabilities

The Eq. 2.18 exhibits the Bayesian theorem for the probability of model parameter values, where $p(\theta|y, M)$ is the probability of the parameter vector θ given the experimental data y and the model M (=posterior probability), $p(y|\theta, M)$ is the probability of the data y given the model M and the parameter vector θ (=likelihood), $p(\theta|M)$ is the probability of the parameter vector θ given the model M (=prior probability) and $p(y|M)$ is the probability of the data y given the model M (=marginal likelihood). Bayesian model parameterization, fit of models respectively, is achieved by means of posterior probabilities of model parameter vectors, which represent all possible model parameterizations (parameter values respectively).

The marginal likelihood is the integral of the numerator of Eq. 2.18 over the whole parameter range (Eq. 2.19).

$$\mathbf{p(y|M)} = \int_{\boldsymbol{\theta}} \mathbf{p(y|\boldsymbol{\theta}, M) \cdot p(\boldsymbol{\theta}|M) \, d\boldsymbol{\theta}}$$

Equation 2.19: Bayesian marginal likelihood

The prior probability of a certain parameter vector is based on prior knowledge about process parameters (e.g. information on parameterization from the literature). If there is no prior knowledge all parameter vectors within the range considered are given the same probability (uniform probability distribution).

The likelihood can be explained as the probability that the measured data occurs, with respect to the error of experimentation, if the model M with the parameterization θ was true. It can be calculated by Eq. 2.20 if a Gaussian probability distribution of the error of experimentation was assumed (as in this work).

$$p(y|\theta, M) = \frac{1}{\sqrt{(2 \cdot \pi)^d}} \cdot e^{-\frac{1}{2} \cdot (y - \mu)^T \cdot \Sigma^{-1} \cdot (y - \mu)}$$

Equation 2.20: Bayesian likelihood (assumed Gaussian distribution)

In Eq. 2.20 the parameter d is the dimension of the parameter space, µ the solution of the model M at parameterization θ and Σ the covariance matrix of the error of experimentation when reproducing y. The latter can be experimentally determined by at least double realization of one defined experiment.

2.3.2.2.2 Posterior probability estimation

A mayor problem of Bayesian analysis is to provide an estimate of the marginal likelihood as this integral is analytically almost never solvable (Chib, S. and Jeliazkov, I., 2001, pp. 270-281).

In determination of the posterior parameter distribution this problem can be avoided by direct sampling from this distribution by Gibbs-sampling which is a special case of the Metropolis-Hastings-Algorithm (Markov-Chain-Monte-Carlo method). It can be shown that this algorithm produces Markov-Chains that converge towards the desired posterior parameter distribution, which means that at chain convergence all following chain steps form a representative sample of the posterior parameter distribution. As this algorithm is based on drawing random numbers from certain probability distributions, it represents a Markov-Chain-Monte-Carlo (MCMC) method.

The approach starts with an arbitrarily chosen initial parameter vector θ from the parameter space considered. This is the first state of the Markov-Chain. Then a new parameter vector, which is the so called "candidate vector" θ^* is drawn randomly from a certain parameter distribution over the parameter space considered. This distribution - by which the candidate vector is proposed - is the so-called "proposal density". The relative posterior probability of current state θ and candidate θ^* can be calculated according to Eq. 2.21. Both posterior probabilities share the same value for the marginal likelihood. The latter can therefore be eliminated from the ratio. Thus the posterior probability ratio can be calculated from the likelihoods and prior probabilities. This ratio is then subject to the probability of move $\alpha(\theta, \theta^*|y, M)$ exposed in Eq. 2.21 which is the probability that the candidate is accepted as next state of the Markov-Chain.

$$\alpha(\theta, \theta^*|y, M) = \min\left\{1, \frac{p(y|\theta^*, M) \cdot p(\theta^*|M)}{p(y|\theta, M) \cdot p(\theta|M)}\right\}$$

Equation 2.21: Probability of move

Eq. 2.21 specifies that θ^* is always accepted if the ratio was larger than 1. Otherwise the acceptance probability equals the ratio value. In practical programming the latter is realized by randomly drawing a random number from a standard uniform distribution. If the random number was smaller than the posterior probability ratio the candidate is accepted, otherwise rejected. Mathematical evidence has been provided that the chain of accepted parameter vectors converges towards the desired posterior probability distribution which means that at chain convergence all following accepted parameter vectors are a representative sample of that distribution. From that sample, the distribution as well as its characteristic constants (expected value, variance) can be estimated. (Chib, S. and Jeliazkov, I., 2001, pp. 270-281; Chib and Greenberg 1995; Hastings, 1970; Tierney, 1994; Hartmann et al., 1974; Kreyszig, E., 1972). In practice Bayesian point estimation, which is the estimation of model parameter values and events (e.g. process outcomes at defined process settings) by Bayesian means, applies the posterior probability distribution (e.g. for model parameters or process outcomes). Characteristic distribution parameters serve as point estimators. The type of point estimator

applied depends on the chosen risk function to be minimized. Risk functions represent different ways of measuring the distance between the estimate and the unknown actual parameter or event. The most common point estimate is the expected value which minimizes the mean squared error of estimation. The expected value of a distribution can be estimated by the arithmetic mean of events, which converges to the expected value as the number of events (sample size) approaches infinity. Another point estimator is the posterior mode which maximizes the probability of matching the actual value (minimizes the uncertainty of estimation). Both point estimators are applied in this work and compared in results and performance. Also other point estimators exist and are occasionally used.

2.3.2.2.3 Müller-Algorithm

The Gibbs-sampling, which is applied in this work to generate a representative sample from the Bayesian posterior parameter distribution, is a version of the Metropolis-Hastings-algorithm, which in turn is a Markov-Chain-Monte-Carlo method. Thus, the representative sample of the posterior distribution is taken as sample from the converged Markov-Chains. The velocity of chain convergence significantly determines the computational effort and duration of this approach. Müller (1991) provided mathematical evidence that the chain convergences is the faster the more similar the proposal density is to the desired posterior distribution. As the chain actually converges towards the posterior distribution an update for the proposal density can be calculated in defined intervals from the Markov-Chain output which then accelerates the chain convergence. The current chain state θ is applied as mean of the proposal density. Müller et al. (1991) state, that the variance of the proposal density is of general importance for the velocity of chain convergence. If the variance is too small, the proposals are on average too close to the current state and thus exhibit a high probability to be accepted. This means that the parameter range to be investigated is not properly screened, which therefore lowers the informative value of the sample. If the variance is too large, the whole parameter range is properly screened but proposals occur more frequently that differ much from the current state and, thus, exhibit a lower probability to be accepted. Due to that fact the variance of the proposal density has to be balanced in a defined interval to realize an acceptance rate that ensures proper screening of the parameter range and chain convergence velocity. Based on practical experience Müller et al. (Müller et al., 1991) define an optimal acceptance rate between 80 % and 20 % and propose a multiplication of the variance by factor 1,2 and 0,7 respectively if this range is exceeded. These tunes of the variance are achieved in a defined interval between the re-estimations of the proposal density from the Markov-Chain.

To limit the computational effort and to gain a representative sample from the posterior distribution it is necessary to identify the step at which the Markov-Chain has converged. Also the number of following Markov-Chain steps that form an appropriate representative sample has to be known. For this purpose a Monte-Carlo simulation is achieved previously to identify the sample size that is required to estimate the distribution parameters (mean, variance) with the desired accuracy. An equal number of identical Markov-Chains is then run in parallel so that the exact state of chain convergence can be determined at every chain point by calculating the mean and variance over all parallel chains at this point. From the point of chain convergence a sample of same size is then taken over all chains as representative sample of the posterior distribution. From that sample, the probabilities for distribution intervals (refered to the interval mean), the expected value and variance (uncertainty in model parameterization) can be estimated for the posterior parameter distributions (Müller et al., 1991). This sample also serves as basis for model discrimination, validation and evaluation of model predictions as exposed in the subsequent sub-chapters for the Bayesian approaches applied here.

2.3.2.2.4 Bayesian model discrimination by Chib and Jeliazkov-method

If alternative models are available, Bayesian analysis provides a probability measure by which these models can be compared and selected. Concomitant with Bayesian theory the probability of a model is given by Eq. 2.22, where M_l is the lth of all available alternative models.

$$p(M_l|y) = \frac{p(y|M_l) \cdot p(M_l)}{\int_M p(y|M) \cdot p(M) \cdot dM}$$

Equation 2.22: Probability of a model

As the denominator in Eq. 2.22 is a constant being equal for each model, the probability ratio of a model M_i and M_j can be calculated according to Eq. 2.23.

$$\frac{p(M_i|y)}{p(M_j|y)} = \frac{p(y|M_i) \cdot p(M_i)}{p(y|M_j) \cdot p(M_j)}$$

Equation 2.23: Probability ratio of a model M_i and M_j

The left side of Eq. 2.23 is also called "posterior odds ratio" where $\frac{p(M_i)}{p(M_j)}$ is the ratio of prior model probability (based on prior knowledge) and $\frac{p(y|M_i)}{p(y|M_j)}$ is the so called "Bayes-factor" (Congdon, P., 2006). According to Eq. 2.19, $p(y|M)$ is equal to the marginal likelihood.

Thus the problem of relative model probability estimation turns out to be a problem of marginal likelihood estimation.

The marginal likelihood is the normalizing constant of Eq. 2.18 which is equal for each parameter vector θ. Thus Eq. 2.18 can be transformed for a defined parameter vector θ^* into Eq. 2.24 that provides an estimate for the marginal likelihood regarding a defined model M_l.

$$p(y|M_l) = \frac{p(y|\theta^*, M_l) \cdot p(\theta^*|M_l)}{p(\theta^*|y, M_l)}$$

Equation 2.24: Estimate for the marginal likelihood - Version 1

The likelihood and prior probability of θ^* can be determined as exposed above. Thus the estimation of the marginal likelihood reduces to the problem of estimating the posterior probability of θ^* which is referred to as posterior ordinate. For practical reasons Eq. 2.24 is transformed to Eq. 2.25 by taking logarithms which is especially important for higher dimensional parameterization problems.

$$\log p(y|M_l) = \log p(y|\theta^*, M_l) + \log p(\theta^*|M_l) - \log p(\theta^*|y, M_l)$$

Equation 2.25: Estimate for the marginal likelihood - Version 2

Chib and Jeliazkov propose a Markov-Chain-Monte-Carlo (MCMC) approach by which the marginal likelihood can be determined by a one block sampling from the posterior parameter distribution (Chib, S. and Jeliazkov, I., 2001, pp. 270-281). The generation of a representative sample of the posterior parameter distribution has been exposed above. It is mathematically proven, that at chain convergence the chain reversibility condition, given by Eq. 2.26, is valid for a current chain state θ and a candidate θ^*.

$$p(\theta, \theta^*|y, M_l) \cdot P(\theta|y, M_l) = p(\theta^*, \theta|y, M_l) \cdot P(\theta^*|y, M_l)$$

Equation 2.26: Chain reversibility condition

Here $P(\theta|y, M_l)$ is the posterior probability of parameter vector θ. $p(\theta, \theta^*|y, M_l)$ is the probability of move from θ to θ^*, which is the product of the probability of proposal $q(\theta, \theta^*|y, M_l)$ and probability of acceptance $\alpha(\theta, \theta^*|y, M_l)$ of θ^* given θ (Eq. 2.27).

$$p(\theta, \theta^*|y, M_l) = q(\theta, \theta^*|y, M_l) \cdot \alpha(\theta, \theta^*|y, M_l)$$

Equation 2.27: Probability of move from $\boldsymbol{\theta}$ m to $\boldsymbol{\theta}^*$

An integration of both sides of Eq. 2.26 over the whole parameter space yields a solution for the posterior probability $P(\theta^*|y, M_l)$ of a candidate θ^*, which is given by Eq. 2.28.

$$P(\theta^*|y, M_l) = \frac{p(\theta, \theta^*|y, M_l) \cdot p(\theta|y, M_l)}{p(\theta^*, \theta|y, M_l)} = \frac{\int_\theta p(\theta, \theta^*|y, M_l) \cdot p(\theta|y, M_l) \, d\theta}{\int_\theta p(\theta^*, \theta|y, M_l) \, d\theta}$$

Equation 2.28: Posterior probability of a candidate $\boldsymbol{\theta}^*$

As these integrals are in most cases analytically not solvable, a solution has to be estimated to gain an estimate of $P(\theta^*|y, M_l)$. Chib and Jeliazkov propose an approach to estimate these integrals from the data output of the Gibbs-sampling (Metropolis-Hastings-algorithm respectively) as exposed above for the generation of Markov-Chains (Chib, S. and Jeliazkov, I., 2001, pp. 270-281). This approach is given by the following Eq. 2.29.

$$\widehat{P}(\theta^*|y, M_l) = \frac{\frac{1}{K} \cdot \sum_{g=1}^{K} q(\theta^{(g)}, \theta^*|y, M_l) \cdot \alpha(\theta^{(g)}, \theta^*|y, M_l)}{\frac{1}{J} \cdot \sum_{j=1}^{J} \alpha(\theta^*, \theta^{(j)}|y, M_l)}$$

Equation 2.29: Posterior probability estimate of a candidate $\boldsymbol{\theta}^*$

Here $\{\theta_1, \theta_2, \dots, \theta_K\}$ is a random sample from the posterior parameter distribution which is produced by the Markov-Chains according to the Gibbs-sampling. $q(\theta^{(g)}, \theta^*|y, M_l)$ is the probability for a proposal of a candidate vector θ^* given a parameter vector $\theta^{(g)}$. $\alpha(\theta^{(g)}, \theta^*|y, M_l)$ is the probability of acceptance of θ^* given $\theta^{(g)}$. The likelihood values which are necessary to calculate the proposal and acceptance probabilities in Eq. 2.29 have to be stored for each element of the posterior parameter sample during the random walk of each Markov-Chain. Chib and Jeliazkov state that a suitable candidate θ^* should have a high probability under the posterior distribution (Chib, S. and Jeliazkov, I., 2001, pp. 270-281). Thus the mode of the posterior distribution that can be estimated from the posterior sample is applied as candidate θ^* here. The sample size has to be chosen high enough so that the sample represents the posterior parameter distribution with the desired accuracy. This is done as exposed above. $\{\theta_1, \theta_2, \dots, \theta_J\}$ is a random-sample from the proposal density $q(\theta^*, \theta|y, M_l)$ with the fixed candidate θ^* as expected value of the assumed Gaussian proposal density. As exposed above the actual distribution of the converging Markov-Chain is applied as proposal density during the Markov-Chain run, as this procedure accelerates chain convergence. From the point at which the converging distribution has reached the posterior distribution its distribution parameters (mean, variance) stay constant.

Therefore the variance of the posterior distribution is applied here for the proposal density. $\alpha(\theta^*, \theta^{(j)}|y, M_l)$ is the probability of accepting $\theta^{(j)}$ from the proposal density sample given θ^*. In practice J is chosen equal to K (Chib, S. and Jeliazkov, I., 2001, pp. 270-281). In the implementation of the methods exposed above in MATLAB all probabilities that are assumed to be Gaussian distributed were normalized by transformation to the standard Gaussian distribution to ensure comparability and interpretability of probabilities and results. To avoid computational problems with higher dimensional probability distribution (as occurred in this work) only decimal logarithms of probabilities were applied. Consequently in practical programming application all equations which are computationally implemented were transformed by the laws of logarithmic calculation. As the method applied here is a Markov-

Chain-Monte-Carlo method, also proposals can occur by chance, which exhibit such low probabilities that they are automatically set to zero by MATLAB and cannot be further mathematically processed, thus causing a process breakdown. This problem is considered computationally by setting such zeros equal to the smallest possible value MATLAB can handle. As such events are very rare and exhibit comparatively extremely low probabilities its contribution to overall samples and probabilities can be neglected. Thus it is justified to set an extremely small value, which is computationally not manageable, to another higher value that is manageable but still comparatively extremely small.

2.3.2.2.5 Bayesian model validation and performance evaluation by Geweke-method

The Bayesian approaches exposed so far account for model parameterization and model discrimination. In that context, a Baysian point estimate for the model parameters is determined, as well as a relative model probability estimate. The model-based estimation of process outcomes, as well as evaluation of simulation results, model performance and model validity, by Bayesian means (point-estimation, uncertainty measures) has not been taken into consideration so far. This important modelling problem is treated by means of prior and posterior probability distributions of simulation results and functions of these results, which are compared to the experimental data.

For model validation, Box introduced the concept to apply checking functions $z_i = g_i(y)$ where z_i is the ith value that is calculated from the data y by the ith function g_i (Box, 1980; Box and Wilson, 1951, pp. 1-45). Geweke (2007) proposed a method to determine the prior and posterior probability distributions of potential experimental data y (process outcome at defined process variable settings respectively) and also defined function values z_i that actually occur if the model tested is true. To provide evidence for proper model performance and validity, the real experimental data y^0 and the function value $z_i^0 = g_i(y^0)$ have to be in support of these distributions. The probability distributions expose, which event could be theoretically observed with respect to the error variance of experimentation if the model was true and thus offer a measure for the uncertainty of model predictions, events to be expected and conclusions drawn from experimental observations.

Eq. 2.30 gives the prior probability distribution of hypothetical events, theoretically measured data y respectively, that could theoretically occur if the model was true.

$$p(y|M_l) = \int_\theta p(\theta|M_l) \cdot p(y|\theta, M_l)d\theta$$

Equation 2.30: Prior probability distribution of hypothetical events

$p(\theta|M_l)$ is the prior probability of the parameter vector θ and $p(y|\theta, M_l)$ the likelihood of the data y regarding θ. It has to be recalled here, that y is a vector in which each element

represents a value for a specific process variable setting (e.g. time point of a time depending process). Thus $p(y|M_l)$ is multi-dimensional in case of multiple process variable settings to be considered. In practice the distribution $p(y|M_l)$ for each single dimension or process variable setting is estimated separately. This is done by random sampling of parameter vectors from the prior parameter distribution $p(\theta|M_l)$. For each parameter vector the model-solution over the whole process variable range considered is calculated from the model M_l. These model solutions are then stored yielding a large sample of solutions for each process variable setting. For each of these solutions of the process variable setting actually analyzed, the probability of theoretically occurring experimental data y is calculated over the whole spectrum of y considered. In other words, for each parameter and process variable setting the likelihood is calculated over the whole spectrum of data y that could theoretically occur in practice. Practically, for each process variable setting the spectrum of data y that is considered is divided into small intervals. Then for each parameter vector of the sample the likelihoods of the interval means are calculated and saved to the referring interval. Subsequently from these probability measures (which are saved for each interval) a probability distribution $p(y|M_l)$ over the data space considered is calculated. This is done for each single process variable setting. The distribution $p(y|M_l)$ for a defined variable setting therefore shows the probability of experimental data to occur if the model was true (with respect to prior knowledge about model parameterization and the error variance of experimentation). The actually experimentally measured data y^0 has to be in support of that distribution. Otherwise the model can be considered as being not suitable to simulate the true underlying process at the referring process setting.

In this work a uniform distribution over the parameter range investigated is chosen as prior knowledge from which a sample is drawn by application of a MATLAB routine that generates uniformly distributed random numbers on an interval [0 1]. The sample size has to be chosen in order to properly represent the distribution. This is done by a Monte-Carlo simulation as exposed above.

$$p(z_i|M_l) = \int_\theta p(\theta|M_l) \cdot p(z_i|\theta, M_l)d\theta$$

Equation 2.31: Prior probability of values z_i

Eq. 2.31 gives the prior probability of values $z_i = g_i(y)$. The meaning of the probabilities in Eq. 2.31 as well as the calculation of the prior probability distribution $p(z_i|M_l)$ is analogues to that of y, but only with the difference that z_i is a function of (calculated from) y. For model suitability and validity the value z_i^0 calculated from the actually experimentally measured data y^0 has to be in support of that prior distribution. Eq. 2.32 and Eq. 2.33 give the posterior probabilities of the data y and the function value $z_i = g_i(y)$, respectively.

$$p(y|y_0, M_l) = \int_\theta p(\theta|y_0, M_l) \cdot p(y|\theta, M_l)d\theta$$

Equation 2.32: Posterior probability of hypothetical events

$p(\theta|y_0, M_l)$ is the posterior probability of parameter vector θ and $p(y|\theta, M_l)$ the likelihood of data y regarding the parameter vector θ.

$$p(z_i|y_0, M_l) = \int_\theta p(\theta|y_0, M_l) \cdot p(z_i|\theta, M_l)d\theta$$

Equation 2.33: Posterior probability of values z_i

$p(\theta|y_0, M_l)$ is the posterior probability of parameter vector θ and $p(z_i|\theta, M_l)$ is the likelihood of z_i regarding the parameter vector θ.

3 Objectives

The overall objective of this work is the model based analysis and optimization of the SPD process towards 3-MCPD and related substances, as well as major oil quality parameters. For this purpose an initial theoretical analysis of the SPD process is required in order to define the major oil quality parameters and to identify the most significant influencing process variables. Based on these preliminary considerations a coherent and sophisticated stochastic process modeling and analysis concept shall be implemented and applied in order to optimize the process variables towards the target values. This concept comprises the RSM for identification of statistically significant linear model terms (effects and interactions of process variables towards the target values) and a Bayesian approach towards model fit, discrimination and analysis. Each method requires an initial implementation and validation by a literature example in order to provide a proper understanding of the method dynamics, as well as a functioning algorithm for a successful application to the real-life problem of SPD optimization. For each target value, linear process models shall be derived and analyzed. A model based process optimization shall be achieved for each single target value and additionally towards multiple optimization goals. The simulated optimum process settings shall be experimentally tested and verified.

3.1 Implementation and application of RSM methodology

The RSM is a statistical approach that provides linear process models based on SDoE. Dependend on the size and setup of the SDoE, process models of different complexity can be derived that comprise diverse effects and interactions of the process variables towards the target values. The SDoE maximizes the information on the true underlying process, which is contained in the experimental data, by a minimum of well designed experiments (control of costs). A more complex SDoE theoretically yields the more precise model by consideration of additional effects. The RSM is applied in this work, as it is a sophisticated non-mechanistic modeling approach, that enables precise modeling of processes without deeper knowledge of the physico-chemical and technological sub-processes. It just quantifies the process output (target values) as function of the process input (process variables) at an optimum cost-value ratio by application of SDoE. In this work a full factorial and orthogonal central composite design shall be setup with respect to the most significant process variables and target values identified before for the SPD. For each target value a model of 1st and 2nd grade with statistically significant model terms (effects and interactions respectively) shall be derived. This shall be initially achieved for a literature example to ensure a proper method implemention and understanding of method dynamics for an efficient application to SPD.

3.1.1 Implementation and validation of RSM principles by literature example

The aim of implementing a simple literature example is to prove that the methods to be applied to modeling of SPD are correctly understood and implemented by algorithms in Matlab. The literature example just provides the initial process data, the mathematical theory and the modeling results. It does not provide the implementation of the method by algorithms in Matlab. So if the programmed algorithm provides identical results from the initial process data then this is considered as proof that the method is understood and correctly implemented.

3.1.1.1 Derivation and analysis of a 1st grade model

The theory of full factorial design setup, as well as derivation of statistically significant and validated 1st grade models, shall be implemented and analyzed for a literature example. This example provides the initial data (range of process variable settings and experimental results for the SDoE) and the final 1st grade model. As proof for a correct method understanding and implementation the method theory shall be implemented by an algorithm in Matlab that produces the same 1st grade model from the provided initial data. This prior study shall ensure a proper method application to the real life problem of SPD optimization.

3.1.1.2 Derivation and analysis of a 2nd grade model

The same procedure as exposed in 3.1.1.1 shall be achieved here for the same reasons regarding the setup of an orthogonal central composite design and derivation of a 2nd grade model, respectively. Again a literature example only provides input data and final results. Additionally, also a complex method for the analysis of multivariate 2nd grade models shall be implemented and verified here by a literature example. In case of optimization towards multiple objectives (especially in case of conflicts of objectives) this method enables the visualization and analysis of areas (process variable settings respectively) which provide the same target values (Lin et al., 2011). This method shall be applied to the real life problem of SPD optimization towards multiple objectives if a comparable situation (areas of identical solutions) occurs.

3.1.2 Application of RSM to the real life problem of SPD optimization

Based on the method implementation and verification for a literature example in 3.1.1, a 1st and 2nd grade model shall be derived from a full factorial and orthogonal central composite design here for the SPD. In 3.1.1 the method shall be implemted by algorithm that can be directly applied to the real life problem at hand. In this subchapter initially the SDoEs shall be set up and the corresponding experiments shall be achieved. Afterwards

1st and 2nd grade models with statistically significant model terms (effects and interactions, respectively) shall be derived from the experimental results and validated. Beyond modeling results, also general observations regarding process and experiments shall be exposed and discussed as they might contribute to a proper understanding of SPD dynamics and optimization.

3.1.2.1 Setup of Design of Experiments

Initially a complex theoretical analysis of the SPD and the optimization problem at hand shall be achieved to identify the mayor process variables and target values that shall be subject to the SDoE and process modeling. For each process variable the range of analysis shall be reasonably defined and applied to the algorithm (programmed in 3.1.1) that provides the full factorial and orthogonal central composite design. The experiments shall be achieved and the pre-defined target values shall be measured for each experimental setup of the DoEs.

3.1.2.2 Model derivation for target value 3-MCPD and related substances

From the experimental results of the full factorial and orthogonal central composite design statistically significant models of 1st and 2nd grade for 3-MCPD and related substances shall be derived and validated by statistical means.

3.1.2.3 Model derivation for target value rancimat

From the experimental results of the full factorial and orthogonal central composite design statistically significant models of 1st and 2nd grade for the rancimat value shall be derived and validated by statistical means.

3.1.2.4 Model derivation for target value acid value

From the experimental results of the full factorial and orthogonal central composite design statistically significant models of 1st and 2nd grade for the acid value shall be derived and validated by statistical means.

3.1.2.5 Model derivation for target value tocopherol

From the experimental results of the full factorial and orthogonal central composite design statistically significant models of 1st and 2nd grade for the tocopherol content shall be derived and validated by statistical means.

3.1.2.6 General observations during the trials

Special characteristics and dynamics observed during experimentation shall be exposed and anylzed regarding their importance for a proper understanding and characterization of the SPD process and result interpretation.

3.2 Bayesian model fit, discrimination and performance analysis

The Bayesian methodology applied in this work is a coherent and self-contained approach for model fit, discrimination and analysis by probibablistic means. In contrast to most conventional approaches, which fit the experimental data (e.g. RLS-methods), it explicitly estimates the true underlying process and hypothetical events by consideration of experimentation noise. Additionally, an uncertainty measure of results is provided by probability distributions over the range of hypothetical alternatives. Initially the Bayesian methodology shall be implemented for a literature example to ensure an effecient method implemention by a functioning algorithm in Matlab, as well as a proper understanding of method dynamics and result interpretation. This preliminary studie is cutting-edge for an efficient application to the real-life problem of SPD modelling and optimization. Thus closing, the Bayesian model fit, discrimination and analysis approach shall be applied to the models derived by RSM for the SPD in order to optimize the process variables towards the single target values, as well as towards multiple optimization goals.

3.2.1 Implementation and validation of the Bayesian methodology by the example of a growth model

The Bayesian methodology applied here for model fit, discrimination and analysis shall be initially implemented for a simple growth model example by an algorithm in Matlab. This algorithm shall be constructed in a way to enable its direct application to the SPD just by exchanging models and initial data in the respective m-files. A verification of method and algorithm, as well as an efficient application to SPD, requires an initial comprehensive study of method dynamics and results by a simple example. This study can be subdivided into six central questions which shall be answered below.

3.2.1.1 Does the implemented method generally provide chain convergence and true process recovery?

Does the implemented theory produce a chain convergence from a prior to a posterior distribution?
How far is the method / algorithm able to recover the true model, parameterization and observation from a direct solution of this known model (simulated non-interfered

experimental data) when starting the MCMC-process from a parameterization much different from the real one?

3.2.1.2 What is the influence of prior knowledge?

In which way does the choice of the prior knowledge influences the analytic results and how can incorrect / inadequate prior knowledge superimpose the information about the true underlying process contained in the experimental data?
This analysis is crucial for the right choice of the prior knowledge in the analysis of the Short-Path-Distillation process.

3.2.1.3 How to interpret the output of the Gibbs-sampling?

How has the output of the Gibbs-sampling, the Markov-Chains respectively, to be interpreted in terms of determination of the posterior parameter distribution and model parameterization (model fit respectively)?

3.2.1.4 Is the method capable to recover the true process?

Is the method and algorithm implemented here able to recover the true underlying model and parameterization from a model solution that has been interfered by noise equivalent to that of the real experimental data?

3.2.1.5 How are the results for the real experimental data influenced by the prior variance?

Which results do the method and algorithm implemented here provide for the real experimental data and how does the choice of the prior variance influence these results?

3.2.1.6 Are the results reproducible with acceptable precision?

Are the results of the Bayesian method applied here reproducible with acceptable precision when the method is applied to identical data sets at identical conditions (start values, a priori knowledge etc.)? In other words: Is the method correctly tuned (sample sizes, number of Markov-Chains etc.) in order to provide a sufficient reproducibility?

3.2.2 Application of the Bayesian methodology to the real life problem of SPD optimization

The Bayesian methodology for model fit, discrimination and analysis, as well as the corresponding algorithm, shall be applied to the SPD models for the target values rancimat value, acid value and tocopherol (derived by RSM, compare 3.1.2). An application to models for 3-MCPD-FE and related substances (e.g. G-FE) wil not be

achieved as it will be shown in chapter 5.1.2 that the contaminant concentrations measured for the SDoE are so low, that model derivation by RSM is not possible here. Thus, the SPD is generally considered as suitable to meet the process requirements with respect to contaminant formation. As already mentioned above the method and its implemention shall be analyzed and verified before to enable an efficient application to SPD (compare 3.2.1). For a comprehensive Bayesian analysis and optimization of the SPD, the following results shall be provided for each target value (model respectively) in terms of point estimates and uncertainty measures: Bayesian model fit, discrimination (in case of alternative models), validation, optimum process variable setting and probabability of experimental observations at this settings. Also the Markov-Chain convergence (during model fit) and the model parameter probability distributions shall be exposed and analyzed. Finally the possibilities of process optimization towards multiple goals shall be discussed.

3.2.2.1 Model fit, discrimination, validation and analysis of rancimat models

A 1st grade and 2nd grade SPD process model, which have been derived by RSM for the target value rancimat, shall be fit and discriminated by the Bayesian methodology applied in this work. Also a model validation shall be achieved by Bayesian means, as well as an analysis of experimental observations to be expected at the optimum process setting. An exemplary exposition and discussion of the Markov-Chain concergence shall provide evidence for a proper sampling from the posterior parameter distribution. The latter shall be exemplary compared to the prior distribution to proof the capability of the method to update prior knowledge and to reduce prediction uncertainties by extracting the information on the true underlying process contained in the experimental data.

3.2.2.2 Model fit, discrimination, validation and analysis of acid value models

The same objectives that have been exposed for the rancimat value (compare 3.2.2.1) are also valid here for the acid value.

3.2.2.3 Model fit, discrimination, validation and analysis of tocopherol models

The same objectives that have been exposed for the rancimat value (compare 3.2.2.1) are also valid here for the tocopherol content.

3.2.2.4 Conclusions for the process optimization towards multiple target values

Based on the validated process models and the respective optimum process settings for the single target values, it shall be discussed if a simultaneous optimization (minimization or maximization within the range of analysis) of all target values is possible. If not, the

conflicts in objectives shall be clearly identified and possibilities for acceptable compromises shall be discussed.

4 Materials and Methods

This chapter comprises the exposition of the SPD plant, as well as the characterization of the feed stock and analytic methods. For the SPD, the complete technological set up and the major process variables are exposed. Also the basic physico-chemical processes, which are involved in the deodorization of edible oil by SPD, are explained. The representation of the analytic methods includes the determination of 3-MCPD-FE and related substances, rancimat value, acid value and tocopherol content.

4.1 Short Path Distillation (SPD)

The term 'short path distillation', or SPD, is derived from the short distance between evaporator and condenser surface (see also Figure 4.1 below). In a SPD unit, the vapor is generated on the evaporator surface from the feed oil, while the latter flows down that surface by gravity. The feed oil is agitated and distributed evenly on the entire inner evaporator surface by rotating whippers / rolls. The vapor stream directly passes over from the evaporator to the condenser surface. As the distance between these surfaces is below the mean free path of vapour molecules, it is called "short path". Thus, re-condensation of vapour on the evaporator surface is avoided and the vapor stream is almost instantly condensed on the condenser surface. This also keeps the vapor from building up vapor pressure, which would lower the mass transport rate due to suboptimal concentration gradients of volatile compounds. The technological design of SPD enables extreme vacuums down to 0.001 mbar. Consequently, heat sensitive materials undergo purification at much lower temperatures due to boiling point lowering, which prevents undesired heat induced formation and degradation processes. Finally, it should be mentioned that SPD in industrial scale is a continuous process.

In comparison to standard batch distillation processes, the special features of SPD enable the separation of liquid feed into liquid fractions, which consist of two or more components, under gentle conditions (Frank, 1969, p. 18; Goffic, F. E. and Albers, M., 2002, pp. 292-295). The special features of SPD are as followed:

- Reduced boiling temperature through fine vacuum as low as 10^{-3} mbar (mild conditions)
- Limited heat exposure by reduced residence time
- No static pressure loss due to thin film evaporation

- Efficient heat transfer

- Large surface per volume unit

- Continuous feed agitation / homogenization (surface renewal, avoidance of temperature gradients)

- Short steam path and one-time vaporization; avoidance of gas haze

SPD belongs to the process group of thin-film distillation technologies (UIC, 2007). In this group, distillation from a thin film takes place at fine vacuum. SPD is often used when standard batch distillation is not suitable to separate thermo-sensitive liquid mixtures without thermal deterioration.

The significant lowering of pressure in the SPD process enables a substantial reduction of the required boiling temperatures. If the pressure is reduced by one decimal power, then the boiling temperature will also be reduced by approximately 25°C (Bethge, 1996, pp. 84-87). This enables the SPD unit to separate heat-sensitive components, which are contained in a certain liquid, with significantly reduced thermal deterioration (Krell, 1976, pp. 303; 312).

In the literature, various types of distillation processes are described. The distillation types can be distinguished based upon the specific pressure range applied (Table 4.1). SPD is thereby rated as applying a medium to low vacuum range, which is also called a fine vacuum (Frank, 1969, p. 19).

Distillation	Vacuum	Pressure range *[mbar]*
Thin-film dist.	rough	1013,3 - 1,3
SPD	low & medium	$1,3 - 1,3 \cdot 10^{-3}$
Molecular dist.	high	$1,3 \cdot 10^{-3} - 1,3 \cdot 10^{-5}$

Table 4.1: Vacuum distillation processes differentiated by applied pressure range

The basic design of an evaporator unit with integrated condenser is demonstrated for the SPD in Figure 4.2. It consists of a cylindric double-walled evaporator with an axially arranged condenser unit inside. Both, the condenser and evaporator can be cooled / heated by various media from the inside due to the double-walled design with accessible innerspace. The liquid to be distilled is delivered to the upper inner surface of the cylinder wall. The liquid material then flows downwards. Through the utilization of additional units, e.g. a wiper assembly of rotating rolls, the liquid is evenly distributed on the inner cylinder (evaporator respectively) surface. Thus, the liquid appears as a thin film on the heated inner surface of the evaporator.
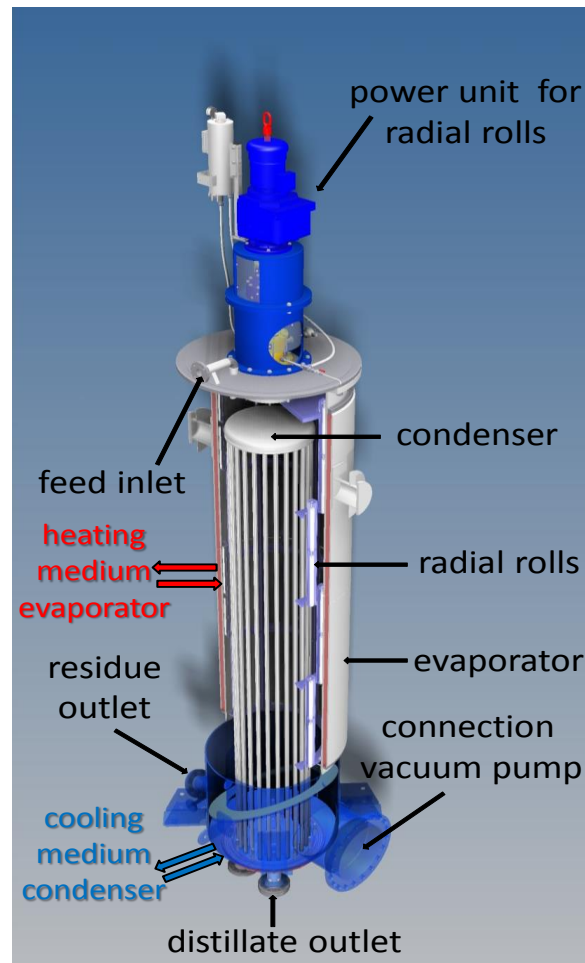
Figure 4.1: Basic design of SPD evaporator unit with integrated condenser

The components with lower boiling point evaporate from the liquid mixture of substances and are afterwards condensed on the surface of the condenser, which exhibits a lower temperature compared to the evaporator. Down the evaporator surface, the feed is separated into distillation residues and distillate, which are finally collected in tempered tanks. From these tanks the liquids can be drained separately. Each volume element of the liquid is less than one minute in direct contact with the heated surface of the evaporator, this is primarily caused by the construction of the SPD plant. Additionally, the liquid is spread as a thin film on the evaporator surface by a stirrer unit, which ensures an optimal heat transfer. Only by this setup a fast heating of the liquid can be achieved (Bethge, 1996, pp. 84-87; Stahl, 1991). Figure 4.2 exhibits the evaporator unit of the technical scale SPD plant VKL-70-5 (VTA GmbH, Niederwinkling) applied for this thesis. The unit is made of glas with an axially arranged spiral condenser inside.

Figure 4.2: Technical scale SPD evaporator unit (glas) with integrated spiral condenser

By axial arrangement of the (e.g. spiral) condenser in the cylindrical evaporator, the distance between evaporator and condenser surface can be minimized. Thus, it is possible to reduce the working pressure of the plant down to $1,3 \cdot 10^{-5}$ mbar. Processing at this pressure range is called molecular distillation. Under these conditions, the distance between the evaporator and condenser surface is below the mean free path of the vapour molecules. This avoids collision of vapour molecules and re-condensation on the liquid film on the evaporator surface. Thus the vapour molecules can directly reach the surface of the condenser, which increases the separation efficiency. In pressure ranges lower than optimal for SPD, the process of diffusion begins. This causes collision of vapour molecules (leaking from the liquid) with diffusing substances in the 'short path', which are on their way from the

47

evaporator to the condenser surface or vice versa (dependent on the actual concentration gradients) (Bethge, 1996, pp. 84-87; Stahl, 1991; Masch, 1950).

The significant dependency of the mean free path of vapour molecules on the pressure is exposed by the example of triglycerides. The mean free path for triglycerides with a molecular weight of roughly 800 g·mol^{-1} is shown for different fine-vacuum pressures in Table 4.3 (Wittka, F. 1940, pp. 557-567).

Pressure *[mbar]*	Mean free path *[mm]*
$10 \cdot 10^{-3}$	7
$4 \cdot 10^{-3}$	25
$1,3 \cdot 10^{-3}$	50

Table 4.2: Mean free path of triglyceride depending on pressure

In a molecular distillation no classical boiling with blistering occurs. Instead, a molecular evaporation, which starts on the surface of the liquid film, can be observed. The equilibrium at the interface between liquid and gaseous phase is disturbed by condensation at the condenser surface and subsequent separation. This equilibrium is repeatedly renewed by the ongoing evaporation process, while the concentration of components with lower boiling point is constantly and significantly reduced in the liquid film (Krell, 1976, p. 322).

With an integrated whipper (e.g. rotating rolls, which are radially arranged directly at the inner evaporator surface, as applied in this thesis), a continuous renewal of the liquid film is achieved.

An n-component mixture is separated during distillation at a constant temperature according to the proportions given by Eq. 4.1 (Hickman, K. C. D.; 1944, pp. 76-77; Krell, 1976, pp. 318-324).

$$\frac{p_1}{\sqrt{M_1}}, \frac{p_2}{\sqrt{M_2}}, \dots, \frac{p_n}{\sqrt{M_n}}$$

Equation 4.1: Separation of n-component mixture at constant temperature

If the temperature increases, the vapor pressure of the involved substances also increases. Thus, a physical correlation exists between these two factors, which is given by the Clausius-Clapeyron equation (Eq. 4.2). Here p is the pressure, T the temperature in Kelvin, ΔH the enthalpy of evaporation and ΔV the change in molar volume at transition from liquid to gaseous phase. Eq. 4.2 can be integrated over the temperature range considered, if the enthalpy of evaporation is considered as being constant over that range. This integration leads over to Eq. 4.3, where R is the gas constant and p_1 and T_1 are the known saturation vapour pressure and referring temperature of an initial state. p_2 is the saturation vapour pressure of interest at temperature T_2. Thus, Eq. 4.3 can be converted into Eq. 4.4 where $\frac{\Delta H}{R}$

and c are constants which are in practice experimentally determined or given by the literature for a defined volatile substance. Given a measured data set of the saturation vapour pressure p at different temperature T, these constants can be experimentally determined as slope and axis intercept from the plot (regression curve respectively) of $\ln p$ versus $\frac{1}{T}$ according to Eq. 4.4. Figure 4.3 examplarly shows the plot of the integrated Clausius-Clapeyron equation (compare Eq. 4.4) for a certain liquid (e.g. fatty acid) with measured data (blue points) and regression curve (dashed blue line). Also a non-logarithmic plot ($p = f(T)$) is common for the vapour pressure curve according to the Clausius-Clapeyron equation.

$$\frac{dp}{dT} = \frac{\Delta H}{\Delta V \cdot T}$$

Equation 4.2: Clausius-Clapeyron equation

$$ln\frac{p_2}{p_1} = \frac{\Delta H}{R} \cdot \left(\frac{1}{T_1} - \frac{1}{T_2}\right)$$

Equation 4.3: Integrated Clausius-Clapeyron equation (version 1)

$$\ln p = \frac{-\Delta H}{R} \cdot \left(\frac{1}{T}\right) + c$$

Equation 4.4: Integrated Clausius-Clapeyron equation (version 2)

The calculation of p at different operation temperature T for volatile oil constituents is, for example, required to identify the ideal process conditions for the distillative separation of substances, as well as for the calculation of the distillate output.

Vapour pressure curves exhibit the dependency of the vapour pressure of defined substances on the temperature and, thus, ease the distillative separation of substances in liquid mixtures as well as the definition of the most suitable process conditions to be applied in this regard. The top peak of the curves indicates the boiling temperature at normal environmental pressure. If the residence time of a volume unit oil in the plant increases, a separation of substances can be reached at a lower temperatures (Hickman, K. C. D. (1944), p. 87; Krell, 1976, p. 334).

Crude palm oil consists of 94%-98% triglycerides (basicly palmitic and oleic acid), 500-700 ppm carotinoids, $60\text{-}100 \ \mathrm{mg} \cdot (100\mathrm{g})^{-1}$ tocopherol/tocotrienol, as well as approx. 5 % diacylglycerids and FFA (O'Brien, 2009, pp. 43–48; Ooi et al., 1994, pp. 423-426).

Undesired oudoring and flavouring substances, which are separated in the final distillation step of refining (the so-called deodorization), are minor components such as FFA, ketones, unsaturated aldehydes, hydrocarbons and lactons (Bockisch, 1993, ch. 7; Ooi et al., 1992).

A separation of these substances is possible because they exhibit higher vapour pressures at defined temperature in comparison to triglycerides, as well as a comparatively smaller molecular weight (Krell, 1976, pp. 318–323).
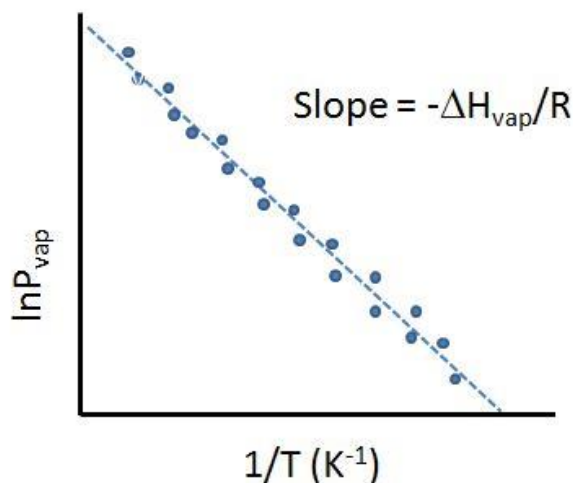
Figure 4.3: Vapor pressure curve of liquid – Plot of integrated Clausius-Clapeyron equation
(Open access source: http://chem.libretexts.org)

A substantial part of the tocopherols is removed from the distillate during the standard deodorization process (Rimbach et al., 2010, ch. 8). Tocopherols are characterized by a molecular weight of $M=415 \text{ g}\cdot\text{mol}^{-1}$ and a saturation vapour pressure of $p=0,2$ mbar (200°C). In comparison to triglycerides, they exhibit a smaller molecular weight but comparable vapour pressure (Fregolente et al., 2006, p. 184).

Various studies provide elimination curves of free fatty acids from palm oil (Rawlings, 1939, pp. 231-232; Holló et al., 1964, pp. 936-941). A maximum temperature of 140°C to 180°C for a complete elimination was determined. Furthermore, it was revealed that triglycerides (molecular weight of $800 \text{ g}\cdot\text{mol}^{-1}$) are distillative separated at temperatures above 200°C. Based on these facts, the definition of the temperature range, to be applied and investigated for the SPD in this work, was achieved.

Although the evaporator unit is the basic feature of SPD, which enables the separation of undesired volatile compounds from the oil by the technological and physico-chemical principles exposed so far above, a complete SPD plant is much more complex. A proper and efficient operation of the evaporator unit requires a lot of technical and technological systems of high complexity. Figure 4.4 and 4.5 exhibit the total setup of the SPD plant VKL-70-5 (VTA GmbH, Niederwinkling) applied for this thesis.
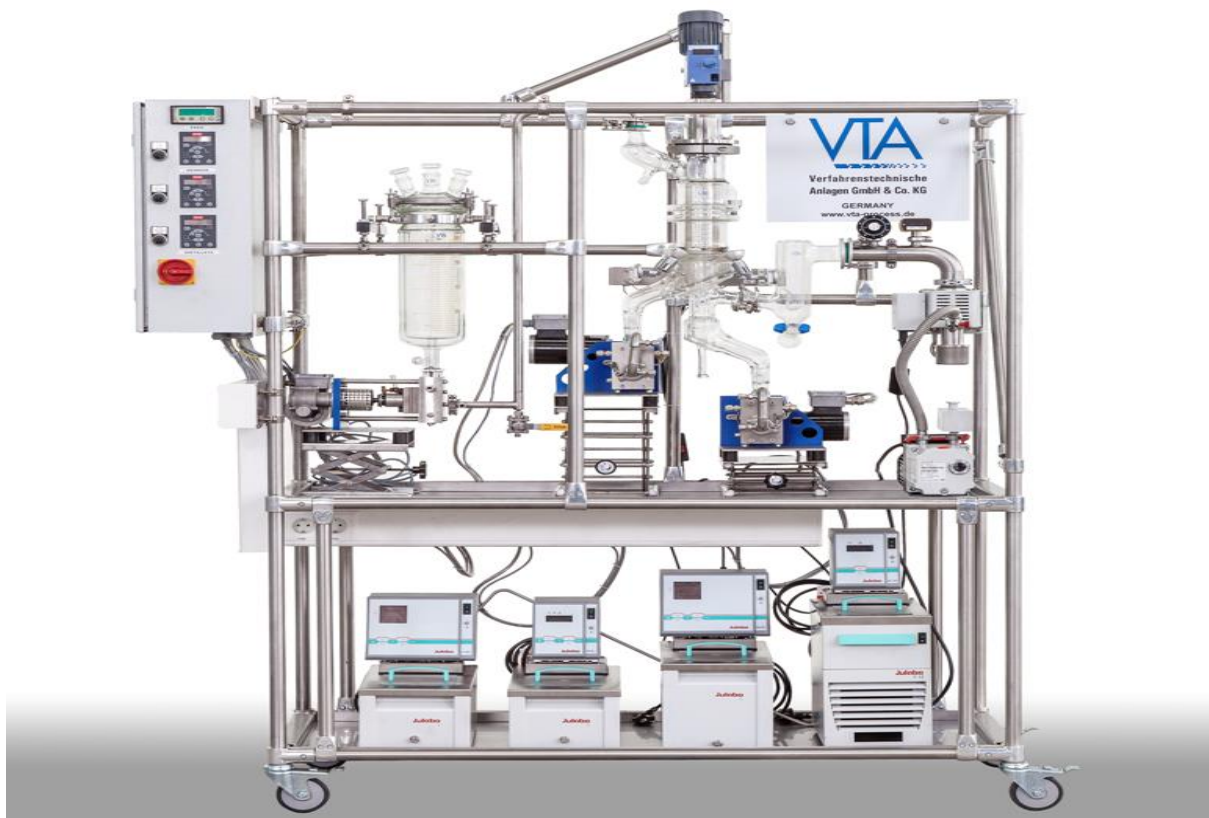
Figure 4.4: Setup of SPD plant VKL-70-5 (VTA GmbH, Niederwinkling)
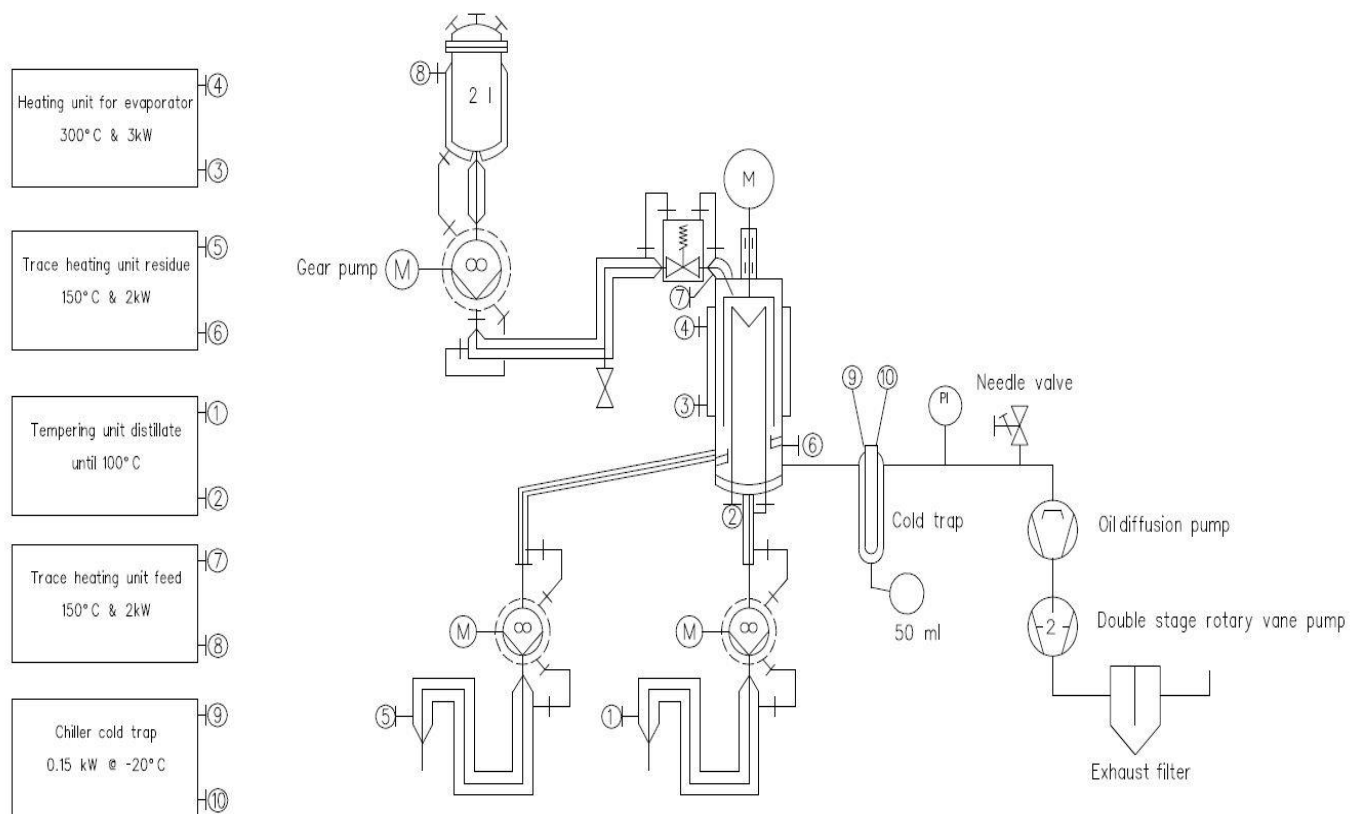
Figure 4.5: Setup scheme of SPD plant VKL-70-5 (VTA GmbH, Niederwinkling)

The following Table 4.4 exhibits the major technical parameters of the SPD plant VKL-70-5 (VTA GmbH, Niederwinkling) applied for this thesis. Annex A comprises all relevant figures and pictures of the SPD process and plant.

Feed supply	
Reservoir	2 L volume, Duran-glas
Pump	Tempered gear pump, 2,8 L·h^{-1} max. feed rate
Pipeline	Stainless steel, double-walled
Evaporator unit VKL70-5 FDRR SKR	
Material	Duran-Glas
Inner evaporator surface	0,052 m²
Max. evaporation temperature	300°C
Outer condenser surface (spiral)	0,080 m²
Whipper system	SKR Block Whipper System (3 whipper strings, PTFE/Graphit)
Whipper speed	40 - 500 rpm
Destillate and residue output	
Pump	Tempered gear pump; 2,8 L·h^{-1} max.feed rate
Evacuation system	
Minimum pressure (valid for clean and dry system)	$< 1 \cdot 10^{-3}$ mbar with diffusion and rotary vane pump $< 1 \cdot 10^{-2}$ mbar with rotary vane pump
Cold trap	2,5 dm²; Duran-Glas
Oil diffusion pump	$40\ L \cdot s^{-1}$ max. feed rate
2-stage rotary vane pump	$2,5\ m³ \cdot h^{-1}$ max. feed rate

Table 4.3: Technical parameters of SPD plant

4.2 Feed stock

All chemicals used in the experiments, unless marked otherwise, were manufactured by Carl Roth GmbH & Co. KG (Karlsruhe, Germany). The bleaching earth was supplied by Süd-Chemie (Munich, Germany).

The crude palm oil used for the experiments was supplied by Cargill B.V. (Schiphol, Netherlands). It was partially refined as part of a previous treatment. The parameters for this standard refining process are further described in Table 4.5. The refined oil was stored at constant cooling (6°C) until it was applied in the SPD experiments.

Prior to the refining process, the crude palm oil showed a tocopherol content of 25,4 $mg \cdot (100g)^{-1}$ and an acid value of 10,3 $(mg\ KOH) \cdot kg^{-1}$. The rancimat value of the raw palm oil was 0,55 h. After the partial refining (compare Tabel 4.5) a comparatively constant acid value of approx. 10,0 $mg\ KOH \cdot kg^{-1}$ was measured; this is equivalent to a FFA content of 4,8 %.

1. Drying	
Temperature *[℃]*	95
Duration *[min]*	15
Pressure *[mbar]*	100
2. Degumming	
Temperature *[℃]*	60 (Acid addition)
	95 (Degumming)
Duration *[min]*	20 (+10 for heating-up)
Pressure *[mbar]*	ca. 990
Citric Acid (50%) *[m-%]*	0.06
3. Bleaching	
Bleaching earth	Tonsil Supreme 118 FF
Temperature *[℃]*	95
Wet-bleaching	
Duration *[min]*	10
Pressure *[mbar]*	ca. 990
Dry-bleaching	
Duration *[min]*	20
Pressure *[mbar]*	20-30
4. Filtration	
Plate filters	
Vaccum filtration	

Table 4.4: Process parameters of the partial refining process of crude palm oil

4.3 Experimental and analytic methods

The methods exposed in this chapter were applied for the realization of the SPD experiments and analysis of the samples.

4.3.1 Experimental methods

The experiments were carried out on a single-stage short path distillation plant (SPD) manufactured by Verfahrenstechnische Anlagen GmbH & Co KG (Niederwinkling, Germany). The specific setup and operation of this type of plant is described in subchapter 4.1.

Precisely, 400 g of bleached palm oil was weighed in a glass beaker for each individual experiment. The substance was then fused in a water bath at a constant temperature of 60°C.

Such refining intermediates often contain varying degrees of gases that are acquired from the surrounding air. For this reason, a degasification process prior to the vacuum distillation is usually a standard step. Moreover, during the course of the injection of the refining intermediates into the empty SPD plant, many soluble gases escape from the liquid. The escaping gases occupy a substantial amount of space within the vacuum, which can initialize foam generation and ultimately lead to a significant increase in the pressure within the SPD plant (Frank, W., 1969, pp. 46–48; Wittka, F. 1940, pp. 557-567; Pingris, A, 1958). No foaming was observed in the experiments of this work. The distillation was performed in four subsequent distillation cycles. Therefore, a separate degasification of the palm oil for each

cycle was waived in the experiments because degasification was completed in the first of the multiple distillation steps.

Initially, the SPD plant was heated to the required temperature and evacuated (p<0,001 mbar) by a rotary vane and a diffusion pump prior to each experiment. Precisely 100 g of the partially refined oil (compare Tabel 4.5) were then placed in the storage tank of the SPD plant, which was kept at a constant temperature of 50°C.

The 100 g (out of 400 g in total) that were placed in the storage tank were used to rinse the unit. The rinsing procedure was performed with exactly the same inflow and agitator speed as required for the next experiment in the SDoE. This step of the process is intended to remove oil residues and contaminants acquired by prior experiments. The distillation residues of this intermediate rinsing step were released from the plant and rejected.

The remaining 300 g of the partially refined oil were then put into the storage tank (at 50°C) and the primary distillation process was initiated with the adequate process parameters defined in the SDoE. The distillation process required constant monitoring in order to avoid the feeding pump from running dry. In such a case, gas could potentially penetrate the plant leading to a reduction of the vacuum.

As four distillation cycles were achieved immediately in succession, the distillation residue of a single cycle was re-introduced into the plant and distilled again. The quality parameter values (e.g. acid value) were only realized through the total separation distance of all cycles. The recommendations for the construction and adjustment of a SPD plant resulting from this thesis mainly refer to that total separation distance.

Upon completion of the four distillation cycles of each individual experiment, the remaining residue was released from the plant and weighed. A condenser temperature of below 60°C resulted in a solid condensate. In this case, the condenser temperature had to be increased in order to re-liquify the distillate and to enable its release from the plant.

Equation 4.5 exposes the correlation of pump frequency v (in *[Hz]*) and volume flow $\dot{V}$ of the feed (in *[ml·min⁻¹]*), which was experimentally determined for the SPD plant applied and for the adjustment range considered in the SDoE. Thus, this correlation is only valid within this adjustment range and aids the derivation of recommendations for up-scaling and design of SPD plants from the results of this work.

$$\dot{V} = 0,6864 \cdot v - 2,5773$$

Equation 4.5: Calculation of feed volume flow from pump frequency

To provide precise data and an error of experimentation for an efficient process modeling and optimization, each experiment was performed twice.

4.3.2 Analytic methods

The acid value was directly determined on-site after each experiment from the remaining distillation residue. The remaining sample was stored in a cool place until it was dispatched to the Max-Rubner-Institute (Detmold, Germany) for further analysis of the other quality parameters.

The following analytic methods were applied for measuring the key quality parameters of the refined and deodorized oil: The total content of 3-MCPD-FE and related substances was determined by the DGF-Method C-VI 18(19) part A and the content of 3-MCPD-FE by the DGF-Method C-VI 18(19), part B. The content of G-FE was calculated by subtracting the content of 3-MCPD-FE from the content of 3-MCPD-FE and related substances. The acid value was determined according to the DGF-Method C-V 2(81), and the Rancimat value with a Rancimat-analyzer manufactured by Metrohm AG. The tocopherol content was determined by the rapid method for the quantitative determination of individual tocopherols in fats and oils (Müller-Mulot et al., 1976., pp. 257-262; Amundsen, N.R., 1966; Hartmann et al., 1974).

The acid value is an indicator of the FFA content in fats and oils. Moreover, this indicator expresses the amount of potassium hydroxide *[mg]* that is needed for the neutralization of the FFA contained in 1 g fat or oil. As this is the only analysis that was achieved self-reliant on-site it is described in detail below.

For each sample, 10 g of the distilled and melted palm oil were dissolved in 40 ml of an ethanol/diethyl ether mixture (1/1 v/v). The FFA contained in the sample was then neutralized by titration with 0,1 N potassium hydroxide solution (normality/factor: 0.966) and neutralization was controled by color changeover of ethanolic thymolphthalein. The acid value was determined three times for each sample. The average value was applied to process modeling and analysis.

The calculation of the acid value (AV) from the amount of potassium hydroxide needed for the neutralization was achieved by use of the following equation:

$$\text{AV} \left[\frac{mg\,KOH}{g\,Fett} \right] = \frac{56{,}1 \cdot \text{F} \cdot \text{a}}{\text{E} \cdot 10}.$$

Equation 4.6: Determination of the Acid Value

The variable a denotes the amount of spent KOH-solution in *[ml]*, F is the normality factor of KOH (here 0,1) and E is the weight in *[g]* of the sample. 56,1 is the molar mass of KOH in *[g·mol^{-1}]*.

5 Analysis and discussion of results

Objective of this work is the model based optimization of SPD towards major oil quality parameters by application of the RSM for model derivation and the Bayesian methodology for model fit, discrimination, validation and analysis. Both approaches are implemented and studied at first for a literature example to ensure a proper functioning and understanding of these methods for the final application to the real-life problem of SPD optimization. Therefore in this chapter the results for method implementation and analysis (example) as well as for the practical method application (SPD) are presented for both, the RSM and the Bayesian approach. As a model to be analyzed has to be derived first, this chapter starts with the RSM (model derivation) followed by the Bayesian methodology (model fit, discrimination, validation and analysis). Both sections are subdivided into the application to an example (implementation, analysis) and the application to SPD (optimization of real-life problem).

5.1 Implementation and application of RSM methodology

Investigations on the effects and interactions of one and more variables towards a target value by SDoE and derived models are known as RSM (Box and Wilson, 1951, pp. 1-45). The following three-dimensional RSM sample graph demonstrates the dependency of the rancimat target value on the input variables pump power (process variable X_4) and evaporator temperature (process variable X_2) for a SPD process with four process variables to be considered. A third statistically significant process variable, the stirrer rotation (process variable X_3), remains constant at optimum setting, as it is not possible to obtain a visual representation of the dependencies of the rancimat value on three variables by a graph. The graph is based on a model of 1st grade derived from a full factorial design.
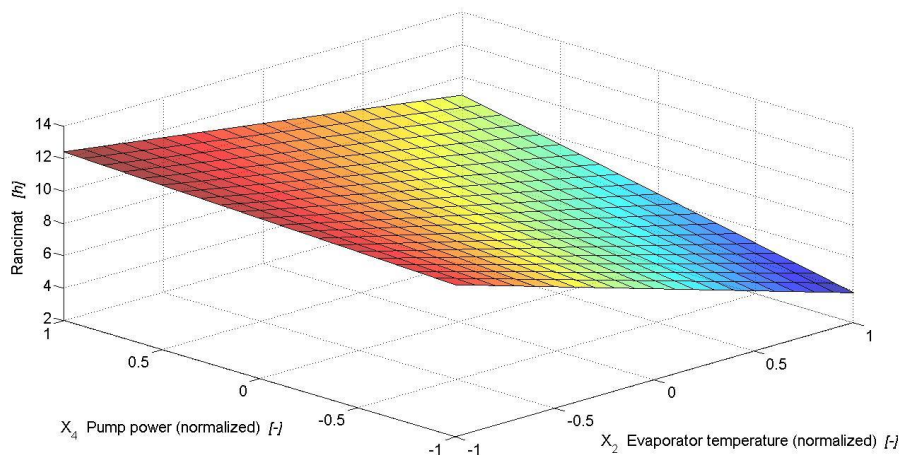
Figure 5.1: Model of 1st grade for rancimat value at optimum stirrer rotation

Models of 1st and 2nd degree derived from full factorial and orthogonal central composite designs provide an approximation of the true underlying process to be optimized towards the

target values. Such approaches are frequently used today, even when little is known about the process itself. The methodology was first introduced in the early 1950s (Davis, O., 1963).

MATLAB is an abbreviation for <u>Mat</u>rix <u>Lab</u>oratory. It is a software that is used by researchers worldwide as tool for technical computing and visualization (modeling, simulation, prototyping, scientific and engineering graphics). The m.file is a MATLAB user script by which mathematical and computational methods / procedures can be implemented and executed by algorithms.

5.1.1 Implementation and validation of RSM principles by a literature example

The aim of implementing this literature example is to proof, that the methods to be applied to modeling and analysis of SPD are correctly understood and implemented by algorithms in Matlab. The example just provides the initial process data, the mathematical theory and the modeling results. It does not provide the implementation of the method by algorithms in Matlab. So if the programmed algorithm provides identical results from the initial process data then this is considered as proof that the method is understood and correctly implemented.

A complex derivation and analysis of a first and second grade model for a literature example (Sundmacher, 2007, pp. 1221-1264) was implemented by the author. This pretest was completed in order to verify, that the RSM method applied here for derivation of a process model for SPD is correctly implemented and thus provides reliable results. The literature example only provided initial process data, some mathematical theory and the modelling results. However, it did not provide sources for the implementation of the method. As the authors' own coding provided identical results from the same process data, the pretest proved that the method could be correctly implemented.

The algorithms / m-files programmed in MATLAB for this investigation will be provided on a storage device and will be available with the thesis. Annex B gives a detailed description of all m-files. The line numbers given in this description refer to the line of code in the respective function.[2]

5.1.1.1 Derivation and analysis of a first grade model

A model of 1st grade is derived here by implementation of an example from the literature, in which the dependency of the tenacity of steel on the carbon, silicium and manganese content is analyzed (Sundmacher, 2007, pp. 1221-1264). The model is derived as exposed in subchapter 2.3.2.1 by application of a full factorial design. The determination of model term

---

[2] The program is provided on a storage device which is part of this work.

significances and model validation is achieved according to the methods applied in the literature example and exposed in chapter 2.3.2.1. Annex B contains a description of the program / algorithm by which this theory has been implemented. The program is provided on the storage device, which is part of that work. This program yields the same results from the same initial conditions as exposed in the literature example. Thus, it can be stated that the theory has been understood and correctly implemented regarding the setup of a full factorial design and derivation of a 1st grade model from its experimental results. Table 5.1 exposes the process variable range, Table 5.2 the full factorial design and referring experimental results and Eq. 5.1 the resulting 1st grade model of the literature example implemented here.

Factor	Range of analysis	
	Min	Max
Carbon *[%]*	0,14	0,26
Silicium *[%]*	0,26	0,56
Manganese *[%]*	0,45	0,55

Table 5.1: Process variable range for literature example RSM 1st grade model

Exp.		Calculation matrix							Tenacity $[N \cdot mm^{-2}]$				
		Trial matrix							Repetition				Mean
	β_0	β_1	β_2	β_3	β_{12}	β_{13}	β_{23}	β_{123}	1	2	3	4	
1	+1	-1	-1	-1	+1	+1	+1	-1	44,4	44,4	45,1	43,8	44,452
2	+1	+1	-1	-1	-1	-1	+1	+1	50,5	51,5	49,2	51,9	50,775
3	+1	-1	+1	-1	-1	+1	-1	+1	48,1	45,4	45,7	48,8	47
4	+1	+1	+1	-1	+1	-1	-1	-1	53,2	54,5	52,1	54	53,45
5	+1	-1	-1	+1	+1	-1	-1	+1	43,3	43,2	46,5	44,7	44,425
6	+1	+1	-1	+1	-1	+1	-1	-1	52,1	49,2	51,9	51,3	51,125
7	+1	-1	+1	+1	-1	-1	+1	-1	48,7	47	50	49,5	48,8
8	+1	+1	+1	+1	+1	+1	+1	+1	53,8	58,8	56,2	56,6	56,35

Table 5.2: Full factorial design and results for literature example RSM 1st grade model

$$\hat{y} = 49,54 + 3,40 * X_1 + 1,86 * X_2 + 0,6313 * X_3 + 0,5438 * X_2 * X_3$$

Equation 5.1: Model 1st grade literature example RSM

5.1.1.2 Derivation and analysis of a second grade model

A model of 2nd grade is derived here by implementation of an literature example, which analysis the dependency of the product yield on the reaction temperature, reaction time and initial molar fraction for a certain chemical production process (Sundmacher, 2007, pp. 1221-1264). The model is derived as exposed in chapter 2.3.2.1 by application of an orthogonal

central composite design. The determination of parameter significances and model validation is achieved according to the methods applied in the literature example and as exposed in subchapter 2.3.2.1. Annex B contains a description of the program / algorithm by which this theory has been implemented. The program is provided on the storage device which is part of that work. This program yields the same results from the same initial conditions as exposed in the literature example. Thus it can be stated that the theory has been understood and correctly implemented regarding the setup of an orthogonal central composite design, as well as the setup and analysis of a 2nd grade model which has been derived from the experimental results. Table 5.3 exposes the process variable range, Table 5.4 the orthogonal central composite design and referring experimental results and Eq. 5.2 the resulting 2nd grade model of the literature example implemented here.

Factor	Range of analysis	
	Min	Max
Reaction temperature *[°C]*	142	152
Initial molar fraction *[%]*	35	40
Reaction time *[h]*	7	10

Table 5.3: Process variable range for literature example RSM 2nd grade model

Exp.	Calculation matrix										Results
		Trial Matrix									
	β_0	β_1	β_2	β_3	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}	Product yield *[%]*
1	+1	-1	-1	-1	1-γ	1-γ	1-γ	+1	+1	+1	55,9
2	+1	+1	-1	-1	1-γ	1-γ	1-γ	-1	-1	+1	70,7
3	+1	-1	+1	-1	1-γ	1-γ	1-γ	-1	+1	-1	67,5
4	+1	+1	+1	-1	1-γ	1-γ	1-γ	+1	-1	-1	68,6
5	+1	-1	-1	+1	1-γ	1-γ	1-γ	+1	-1	-1	63,3
6	+1	+1	-1	+1	1-γ	1-γ	1-γ	-1	+1	-1	68,0
7	+1	-1	+1	+1	1-γ	1-γ	1-γ	-1	-1	+1	68,8
8	+1	+1	+1	+1	1-γ	1-γ	1-γ	+1	+1	+1	62,4
9	+1	δ	0	0	δ^2-γ	-γ	-γ	0	0	0	67,8
10	+1	–δ	0	0	δ^2-γ	-γ	-γ	0	0	0	63,1
11	+1	0	δ	0	-γ	δ^2-γ	-γ	0	0	0	68,4
12	+1	0	–δ	0	-γ	δ^2-γ	-γ	0	0	0	66,2
13	+1	0	0	δ	-γ	-γ	δ^2-γ	0	0	0	68,0
14	+1	0	0	–δ	-γ	-γ	δ^2-γ	0	0	0	65,4
15	+1	0	0	0	-γ	-γ	-γ	0	0	0	67,7

Table 5.4: Orthogonal CCD and results for literature example RSM 2nd grade model

(γ=0,73 ; δ=1,22)

$$\hat{Y} = 67{,}08 + 1{,}81x_1 + 1{,}11x_2 - 1{,}32x_1^2 - 3{,}09x_1x_2 - 2{,}19x_1x_3 - 1{,}21x_2x_3$$

Equation 5.2: Model 2nd grade literature example RSM

For the example taken from the literature the RSM yields an equation of the form

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{11}x_1^2 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3$$

Equation 5.3: Model 2nd grade literature example RSM (general form)

The corresponding response surface of the 2nd grade model cannot be directly visualized and analyzed, as the target value depends on three process variables. To account for that deficit, Eq. 5.3 is transformed into the normal form by coordinate transformation. Eq. 5.3 is given in matrix form by Equation 5.4.

$$\hat{y} = b_0 + b^T x + x^T B x$$

Equation 5.4: Model 2nd grade in matrix form

$$\text{with } b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \text{ and } B = \begin{bmatrix} b_{11} & \frac{b_{12}}{2} & \frac{b_{13}}{2} \\ \frac{b_{12}}{2} & b_{22} & \frac{b_{23}}{2} \\ \frac{b_{13}}{2} & \frac{b_{23}}{2} & b_{33} \end{bmatrix}.$$

By achieving the coordinate transformation

$$(x - x^*) = Mz$$

Equation 5.5: Coordinate transformation

the normal form of Eq. 5.4 is given as

$$\hat{y} - \hat{y}(x^*) = \lambda_1 z_1^2 + \cdots + \lambda_m z_m^2$$

Equation 5.6: Model 2nd grade in normal form

where $(x - x^*)$ is the difference between an coordinate vector x and the extremum coordinate vector x^* (the function value $\hat{y}$ reaches a extremum here), $\hat{y} - \hat{y}(x^*)$ is the difference between the function value $\hat{y}$ and extremum $\hat{y}(x^*)$, M is the matrix of normalized eigenvectors and Λ is the diagonal matrix of the eigenvalues λ of matrix B. This coordinate transformation procedure applied here is referred to as the canonic analysis of the RSM and the resulting Eq. 5.6 is referred to as the Response Surface. The coordinates x^* of the extremum $\hat{y}(x^*)$ are given by Eq. 5.7.

$$x^* = -0{,}5B^{-1}b$$

Equation 5.7: Coordinates of extremum

When $\hat{y}$, $\hat{y} - \hat{y}(x^*)$ respectively, is fixed to a defined value, then Eq. 5.6 can be easily analyzed and visualized in a 3-D plot. In that case the equation describes (the plot shows, respectively) all process settings in z-coordinates that result in the fixed function value, fixed difference respectively. When $\hat{y}$ is chosen equal to $\hat{y}(x^*)$, and thus $\hat{y} - \hat{y}(x^*) = 0$, then the resulting equation shows all process settings at which the extremum is reached. This case is therefore suitable for analysis of process optimization. When the process has to be optimized towards multiple target values than the intersections of the referring optimization equations contain all process settings at which all target values are optimized. In the vector / matrix b and B of the non-transformed Eq. 5.4, for each process variable the effects of that variable on the target value (direct single effects, interactions) are stored as parameter values. Eq. 5.7 enables the calculation of a process setting x^* at which the target value reaches an extremum. This extremum is then applied to the coordinate transformation in Eq. 5.5 leading to the normal form given by Eq. 5.6, which gives the difference between target value and extremum at a defined setting of the transformed process variable z. When the left side of that equation is set to zero then the resulting equation describes all process settings at which the extremum is reached. It is very important to emphasize that it is not known so far, if x^* is a minimum or maximum. Thus process optimization requires further analysis of the Response Surface. This analysis is achieved as explained and visualized (Figure 5.2) below.
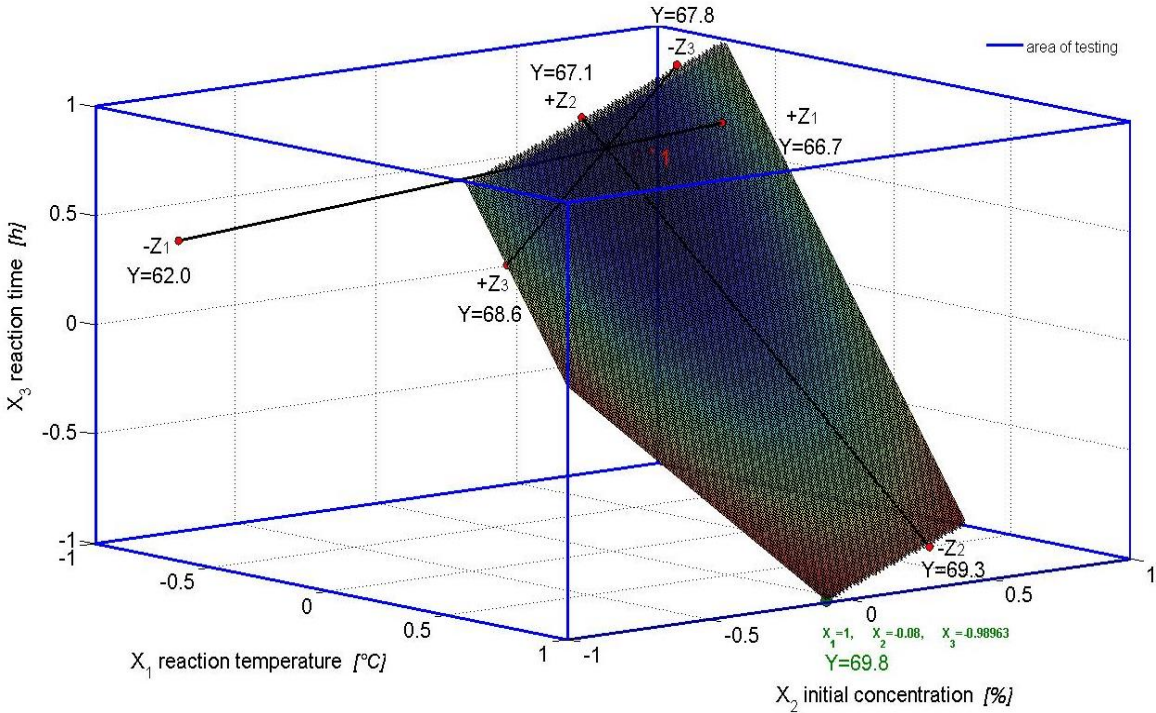
Figure 5.2: Response Surface in x-coordinate system (normalized, non-transformed) for literature example

61

The equation of the Response Surface shows, that one process setting, at which the extremum is reached, is located in the origin of the z-coordinate system where all z-variables are set to zero. At this process setting a yield of 67,08 % is realized. To analyze if the extremum is a minimum or maximum and to understand, how the target value changes with variation of the z-coordinates, the z-axes are plot into the x-coordinate system within the range of x-variables that is subject to investigation. The latter is exposed as cube (room of investigation) in Figure 5.2. Additionally the coordinates of the intersections of the z-axes with the side surfaces of the cube and the referring target values are calculated. This procedure is applied to analyze how the target value changes in the direction of the particular z-axes.
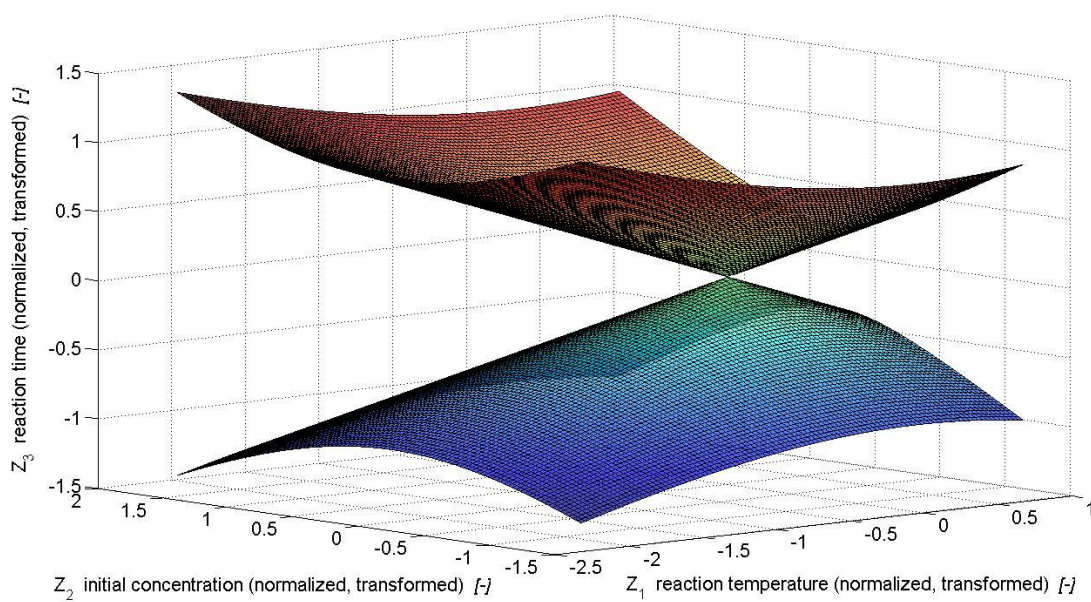
Figure 5.3: Response Surface - Process variable settings for yield extreme value (literature example)

Figure 5.2 shows, that the origin (yield 67,08%) is a minimum and that it is most suitable to realize higher target values by process variable variation in the plain spanned by the Z_2- and Z_3-axis. The contour plot of this plain in Figure 5.2 shows how the target values change with variation of the Z_2- and Z_3-variables (the target value increases with transition of the color from the blue into the red color spectrum). The green point marks the z-coordinate setting at which the maximum for the target value in the Z_2- and Z_3-plain is reached. It has to be outlined here that the plot of the Z_2- and Z_3-plain in the origin just shows the dynamics of target value change in that plain. Nevertheless it is possible that higher target values can be reached within the room of investigation in another parallel Z_2- and Z_3-plain. When a certain target value of interest has been defined, then the left side of the Response Surface function turns to a defined constant value, so that all z-coordinate settings which result in the defined target value can be visualized by a 3-D plot in the z-coordinate system. Figure 5.3 shows

exemplaryly all z-variable settings at which the minimum is reached. Such plots further serve as basis for defining the most suitable process setting for a defined target value with respect to other target values or economical criteria (e.g. lowest possible temperature to reduce energy costs or detrimental heat exposure).

5.1.2 Application of RSM to the real life problem of SPD optimization

In this chapter linear stochastic process models of 1st and 2nd grade are derived for the real life problem of SPD optimization according to the RSM theory exposed in subchapter 2.3.2.1. The implementation of that approach by algorithms in Matlab is exposed by the program description in Annex B.[3] The process models derived here by RSM are further subject to Bayesian model fit, discrimination and analysis in 5.2.2.

5.1.2.1 Setup of Design of Experiments

An efficient process modeling and optimization is based upon a careful setup of a SDoE, which in turn requires a thorough analysis and selection of significant product quality parameters and influencing process variables. This initial analysis and SDoE setup is exposed below.

As target values for the optimization of SPD the following oil quality parameters have been identified: Acid value, rancimat value and tocopherol content, as well as 3-MCPD- and G-FE content.

The acid value is a measure for the FFA content and the rancimat value is a measure for the oxidation stability. Tocopherols support the oxidation stability due to its antioxidant properties and are of nutritional importance as they contribute to the supply of liposoluble vitamins.

For an SPD plant, the following factors have been identified as most influencing regarding the target values: Temperature of the evaporator ($T_{Evaporator}$ *[°C]*), temperature of the condenser ($T_{Condenser}$ *[°C]*), stirring speed ($N_{Stirrer}$ *[rpm]*) of the rotating whipper unit and the feed flow rate of the feeding pump, which is adjusted by the pump frequency (P_{Pump} *[Hz]*). The correlation between pump frequency and feed flow rate is given by Eq. 4.5. The identification of these process parameters and their dependent quality parameters can be considered as a result and a pre-condition of the SDoE. Also the residence time of an oil volume unit in the evaporator unit of the SPD plant, as well as the thickness of the oil film on the evaporator surface, can be considered as significant process variables influencing the oil quality. As these factors are a function of stirrer speed and pump requency anyway, and could not be experimentally determined in this experimental setup, they are not further considered in this work.

To measure and model the impact of these process variables on the quality parameters, an experimental design (SDoE) has to be defined at first. This step must be in accordance with

---

[3] The program is provided on a storage device which is part of this work.

the theoretical principles that are described in Chapter 2.3.2.1. The Response Surface Methodology (RSM) enables the derivation and analysis of process models based on the SDoE results. The investigation ranges of the process variables are primarily dependent on the technical adjustment capabilities of the plant itself. Also the results of pretests and theoretical considerations of the process play an important role in that context.

The temperature of the evaporator was varied between 150°C and 210°C, as pretests showed that temperatures below 150°C could not produce acceptable acid values even when the distillation process was repeated in more than one cycle. When the evaporator temperature was adjusted to the maximum value of 210°C, two cycles of the palm oil through the SPD plant were required in order to achieve an acid value in the predefined range. The second cycle is necessary due to the limited separation range of this type of plant. If the temperature was reduced to the minimum of 150°C, four cycles were required to achieve the same result. In this case, the acid value is reduced in each cycle of the distillation process through a constant deposition of FFA. To provide acceptable acid values and comparability of experimental results for process modeling, in each experiment of the SDoE four cycles of the SPD were achieved.

The choice of the temperature range for the SDoE was also influenced by the literature. However, only experimental assemblies that used the same pressure range for deodorization and de-acidification were analyzed in that context.

Ooi identified an optimal evaporation temperature ($T_{Evaporator}$ *[°C]*) for the process in the range between 150°C and 170°C (Ooi et al., 1996). In comparable experiments other authors found, that for a complete separation of FFA a temperature of at least 180°C is essential. Furthermore, undesirable flavorings and fragrances could successfully be eliminated from the oil at this temperature. These studies also concluded that the separation of triglycerides begins at temperatures above 200°C (Rawlings, 1939; Holló et al., 1964, pp. 936-941; Masch, 1950). These findings were the main reason to fix the maximum temperature to 210°C for the test series.

The expectation before testing was, that with an increasing evaporator temperature, the output of distillate would also increase due to an increased vapor pressure. The latter is caused by the fact that at higher evaporator temperatures the molecules absorb more energy, which is required for the transition from liquid to gaseous phase. In this state, gas can leave the thin oil film spanned on the surface of the evaporator (Frank, 1969, pp. 18; 52-54).

The condenser temperature was varied in the test series between 40°C and 78°C. Thus, the impact of this process variable on the degree of separation could be determined by analysis of the target values in the deodorized oil (residue). One of the experimental results was, that

condensates had a firm consistency at a condenser temperature of below 60°C. At condenser temperatures of 60°C and higher, the condensates had a liquid consistency.

Stirring speed ($N_{Stirrer}$ *[rpm]*) and frequency of feeding pump (P_{Pump} *[Hz]*) were adjusted from the minimum values of 100 rpm and 20 Hz to the maximum values of 390 rpm and 86 Hz. Both process variables influence the thickness of the oil film on the evaporator surface. They also influence the residence time of oil in the evaporator chamber. The ultimate target of adjusting these process variables is to achieve a continuous and homogeneous oil film of minimal thickness being spanned on the evaporator surface. Moreover, the contact time between palm oil and evaporator surface should be long enough to guarantee the separation of undesired volatile substances. This aspect can be fine-tuned with minor adjustments in the feed flow rate.

The principal task of the whipper unit is to achieve a balanced distribution of the liquid film on the surface of the evaporator. The liquid film must have a predefined thickness and it must be constantly renewed to guarantee the homogeneity of the oil film (Frank, 1969, p. 18).

The evaporation efficiency of FFA and tocopherols from oil depends on the consistency and the flow speed of the oil film. The expectation was that the contents of both substances in the oil would be gradually reduced with an increasing stirrer speed and with a decreasing feeding speed. Decreasing the feeding speed makes the oil film thinner and prolongs the time of contact with the evaporator surface. In theory, the separation of volatile FFA, flavorings, fragrances and tocopherols is stimulated by a decreased feeding speed.

It is well-known that rancidity is accelerated by oxidative processes (Metrohm, 2015). Stability and sensory characteristics of oil are negatively correlated with its FFA content and positively correlated with its tocopherols content (Rimbach et al., 2010, pp. 171; 180-181). A high rancimat value is therefore a good indicator of proper oxidation stability and shelf-life.

The temperatures of the feed and distillation residues were kept constantly at 50°C throughout the whole experimental design. The impact of those parameters on the quality of the product is insignificant. Thus they were adjusted to the minimum technologically possible temperature. By this adjustment additional thermal stress for the product can be avoided. With a temperature of 50°C palm oil is still liquid and therefore technologically easy to handle. The pressure was adjusted to its technologically achievable minimum value of 10^{-3} mbar.

The experiments for the deodorization of palm oil by SPD were performed according to an orthogonal CCD with a 2^4-full factorial design as the so-called "kernel" or "core" of the CCD. Based on this SDoE, models with statistically significant model terms (effects and interactions, respectively) were derived according to the theory exposed in chapter 2.3.2.1.

Based on the results of the full factorial design, models of 1st grade were determined. A model of 1st grade describes the dependency of the target value on linear effects of the process variables and their interactions up to highest order.

Furthermore, models of 2nd grade were set up based on the results of the orthogonal CCD. Models of 2nd grade, however, always require a significantly higher number of experiments (cost factor). This in turn provides more information on the SPD process. In theory, such models can therefore more precisely approach the true underlying process by additional identification of non-linear quadratic effects of the process variables towards the target values.

The general approach is to start with a full factorial design to derive models of 1st grade, which demands less testing effort. This approach enables process modeling by identification of linear effects of process variables as well as interactions. Under favorable conditions, this might be sufficient for the model based treatment of the optimization problem at hand. To decide on further ex-ante testing by additional experiments of an orthogonal CCD, a careful weighing of the particular cost-benefit relationship and the required degree of accuracy is neccessary. Through the reduced number of experiments in less complex experimental designs, a reduction of costs and effort is achieved. A potential loss in information is consciously accepted in this case.

To generally evaluate the performance and sufficiency of a model of 1st grade, models of 1st and 2nd grade are set up and compared with respect to their experimental effort, their preciseness and the expected potential benefit of the additional information gained by using a model of 2nd grade (Scheffler, 1986). From these results, proposal can be derived for the optimal choice of experimental designs for modeling and optimization approaches of other comparable and / or scaled-up SPD setups. For example, significant effort and costs can be saved if a 1st grade model is considered as being generally sufficient (in terms of precision and cost-benefit ratio) for modeling and optimization of comparable SPD setups.

Process variable	Non-normalized			Normalized (dimension free)		
	Min.	Max.	Mean	Min.	Max.	Mean
$X_1 = T_{Condenser}$ *[°C]*	40	78	59	-1	1	0
$X_2 = T_{Evaporator}$ *[°C]*	150	210	180	-1	1	0
$X_3 = N_{Stirrer}$ *[rpm]*	100	390	245	-1	1	0
$X_4 = P_{Pump}$ *[Hz]*	20	86	53	-1	1	0

Constant	Adjustment
T_{feed}, $T_{residue}$ *[°C]*	50
Pressure *[mbar]*	10^{-3}

Table 5.5: Process variable range and process constants of SPD for a 2^4 full factorial design

Table 5.5 shows the minimum and maximum values of the key process variables for the 2^4 full factorial design, which is also the kernel of the orthogonal CCD. Such values must be determined ex-ante when the test plan is set up. The experimental settings of a full factorial design are direct combinations of these values. The additional experimental setups for the orthogonal CCD (Table 5.7) have to be calculated with the use of formulas according to the

theory exposed in chapter 2.3.2.1. The matrices of the full factorial and orthogonal central composite design in Table 5.6 and 5.7 provide the normalized process variable settings for experimentation and model term estimation. The conversion between normalized and non-normalized values follows Eq. 2.3 in chapter 2.3.2.1.

experiment	x_0	x_1	x_2	x_3	x_4	$x_1 \cdot x_2$	$x_1 \cdot x_3$	$x_1 \cdot x_4$	$x_2 \cdot x_3$	$x_2 \cdot x_4$	$x_3 \cdot x_4$	$x_1 \cdot x_2 \cdot x_3$	$x_1 \cdot x_2 \cdot x_4$	$x_1 \cdot x_3 \cdot x_4$	$x_2 \cdot x_3 \cdot x_4$	$x_1 \cdot x_2 \cdot x_3 \cdot x_4$
		trial matrix														
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	-1	-1
3	1	1	1	-1	1	1	-1	1	-1	1	-1	-1	1	-1	-1	-1
4	1	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1
5	1	1	-1	1	1	-1	1	1	-1	-1	1	-1	-1	1	-1	-1
6	1	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1
7	1	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1
8	1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1
9	1	-1	1	1	1	-1	-1	-1	1	1	1	-1	-1	-1	1	-1
10	1	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1
11	1	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1
12	1	-1	1	-1	-1	-1	1	1	-1	-1	1	1	1	-1	1	-1
13	1	-1	-1	1	1	1	-1	-1	-1	-1	1	1	1	-1	-1	1
14	1	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1
15	1	-1	-1	-1	1	1	1	-1	1	-1	-1	-1	1	1	1	-1
16	1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1

The top of the table is labelled "calculation matrix".

Table 5.6: 2^4 full factorial design (normalized)

The 2^4 full factorial design is the so called "kernel" or "core" of the corresponding larger orthogonal CCD. The latter additionally exhibits six star points and one central point. A complete orthogonal CCD for four process variables is presented in Table 5.7. The normalized settings of the process variables x_1, x_2, x_3 and x_4 represent the trial matrix which exhibits the actual experimental settings / adjustments. The additional settings of the model constant x_0 and the interactions are not practically adjustable. They result from multiplication of the settings of the corresponding process variables and are applied for parameter estimation of the respective model terms (compare theory in chapter 2.3.2.1). In Tabel 5.7 the interactions of higher than 1st order of the 1st grade model (e.g. $x_1 x_2 x_3$), as exposed in Tabel 5.6, are replaced by quadratic effects (e.g. x_1^2-Y).

exp.	calculation matrix															subset
		trial matrix														
	X_0	X_1	X_2	X_3	X_4	X_1^2-γ	X_2^2-γ	X_3^2-γ	X_4^2-γ	X_1X_2	X_1X_3	X_1X_4	X_2X_3	X_2X_4	X_3X_4	
1	1	1	1	1	1	0,2	0,2	0,2	0,2	1	1	1	1	1	1	
2	1	1	1	1	-1	0,2	0,2	0,2	0,2	1	1	-1	1	-1	-1	
3	1	1	1	-1	1	0,2	0,2	0,2	0,2	1	-1	1	-1	1	-1	
4	1	1	1	-1	-1	0,2	0,2	0,2	0,2	1	-1	-1	-1	-1	1	
5	1	1	-1	1	1	0,2	0,2	0,2	0,2	-1	1	1	-1	-1	1	
6	1	1	-1	1	-1	0,2	0,2	0,2	0,2	-1	1	-1	-1	1	-1	
7	1	1	-1	-1	1	0,2	0,2	0,2	0,2	-1	-1	1	1	-1	-1	core:
8	1	1	-1	-1	-1	0,2	0,2	0,2	0,2	-1	-1	-1	1	1	1	2^4 full
9	1	-1	1	1	1	0,2	0,2	0,2	0,2	-1	-1	-1	1	1	1	factorial
10	1	-1	1	1	-1	0,2	0,2	0,2	0,2	-1	-1	1	1	-1	-1	design
11	1	-1	1	-1	1	0,2	0,2	0,2	0,2	-1	1	-1	-1	1	-1	
12	1	-1	1	-1	-1	0,2	0,2	0,2	0,2	-1	1	1	-1	-1	1	
13	1	-1	-1	1	1	0,2	0,2	0,2	0,2	1	-1	-1	-1	-1	1	
14	1	-1	-1	1	-1	0,2	0,2	0,2	0,2	1	-1	1	-1	1	-1	
15	1	-1	-1	-1	1	0,2	0,2	0,2	0,2	1	1	-1	1	-1	-1	
16	1	-1	-1	-1	-1	0,2	0,2	0,2	0,2	1	1	1	1	1	1	
17	1	1,414	0	0	0	1,2	-0,8	-0,8	-0,8	0	0	0	0	0	0	
18	1	-1,414	0	0	0	1,2	-0,8	-0,8	-0,8	0	0	0	0	0	0	
19	1	0	1,414	0	0	-0,8	1,2	-0,8	-0,8	0	0	0	0	0	0	
20	1	0	-1,414	0	0	-0,8	1,2	-0,8	-0,8	0	0	0	0	0	0	star points
21	1	0	0	1,414	0	-0,8	-0,8	1,2	-0,8	0	0	0	0	0	0	
22	1	0	0	-1,414	0	-0,8	-0,8	1,2	-0,8	0	0	0	0	0	0	
23	1	0	0	0	1,414	-0,8	-0,8	-0,8	1,2	0	0	0	0	0	0	
24	1	0	0	0	-1,414	-0,8	-0,8	-0,8	1,2	0	0	0	0	0	0	
25	1	0	0	0	0	-0,8	-0,8	-0,8	-0,8	0	0	0	0	0	0	center point

Table 5.7: Orthogonal central composite design for four process variables (normalized)

To avoid potential disruptive effects, the experiments were not performed in the numerical order of the experimental design. The true order was determined randomly (Scheffler, 1986, p. 14).

The set up of the SDoE, as well as the derivation of the process models from the results, was achieved by the program described in Annex B according to the theory exposed in Chapter 2.3.2.1.

All experiments of the SDoE were achieved twice with the arithmetic mean being applied to all further calculations such as model derivation. No outliers were defined in order to prove the capability of the Baysian approach to filter noise from the experimental data. The models derived from the SDoE results enable the calculation of optimum process settings.

In the following subchapters all presented values for experimental data and model parameters are rounded up to the second decimal place. In practical programming and computing all calculations were achieved with the highest possible accuracy.

5.1.2.2 Model derivation for target value 3-MCPD and related substances

Experiment	1	2	3	4	5	6	7	8	9	10	11	12	13
3-MCPD and related substances *[mg*kg-1]*	0,13	0,15	0,53	0,05	0,40	0,23	0,10	0,18	0,13	0,08	0,05	0,18	0,18
Experiment	14	15	16	17	18	19	20	21	22	23	24	25	
3-MCPD and related substances *[mg*kg-1]*	0,43	0,15	0,13	0,15	0,10	0,15	0,25	0,43	0,08	0,13	0,08	0,18	

Table 5.8: Results orthogonal central composite design for 3-MCPD and related substances

Table 5.8 exhibits the results of the orthogonal CCD for 3-MCPD and related substances. The measured values are in all cases very close to or already within the standard deviation of experimentation. Thus a model derivation is not possible here as a statistically significant contaminant formation, which can be clearly differentiated from the error in experimentation / measurement, cannot be observed here. In either case, the measured values are close to zero and far below all discussed threshold values and critical concentrations. The formation of 3-MCPD and related substances in SPD is therefore negligible and not subject to further analysis and modeling. It can be stated here, that the SPD process applied in this work is suitable to produce nearly contaminant free deodorized oils within the analyzed range of process variable settings. Thus regarding contaminant formation, the SPD is a promising alternative and sophisticated technological approach for a sustainable production of healthy refined edible oils.

5.1.2.3 Model derivation for target value rancimat

Table 5.9 shows the results for two repetitions of the orthogonal CCD for the rancimat value. The model derivation has been achieved on the mean values and without definition and exclusion of outliers. This procedure should provide evidence that the RSM and Bayesian approach applied here is capable of handling even experimental data that is significantly interfered by noise.

Experiment		1	2	3	4	5	6	7	8	9	10	11	12	13
Rancimat *[h]*	Run 1	8,68	3,51	9,6	3,7	12,4	12,2	12,3	11,6	7,72	3,88	10,2	3,64	12,3
	Run 2	8,77	4,02	9,96	4,09	13,3	11,4	12,7	11,4	7,25	3,62	8,4	4,07	11,2
Experiment		14	15	16	17	18	19	20	21	22	23	24	25	
Rancimat *[h]*	Run 1	11	12,5	11	10,2	10	4,4	11,3	9,42	10,1	10,7	3,97	10,8	
	Run 2	11	12,5	11,3	9,95	9,4	3,96	12,7	10,1	9,75	11,8	3,55	9,15	

Table 5.9: Results orthogonal central composite design for rancimat value

Table 5.10 exhibits the derivation of 1st and 2nd grade process models with significant model terms, which have been determined by six different methods of statistical significance testing. The model derivation and validation has been achieved according to the methods exposed

theoretically in chapter 2.3.2.1. It can be observed that these methods yield four different process models with statistically significant model terms. None of these models proofed to be valid according to the statistical validation method and criterion applied here. The Tables 5.10., 5.12 and 5.14 expose the statistically significant model parameters ß, model terms resoectively, for the target values rancimat value, acid value and tocopherol according to Eq. 2.2 and Eq. 2.10. $ß_0$ refers to the model constant, $ß_1$ to the linear effect of the condenser temperature, $ß_2$ to the linear effect of the evaporator temperature, $ß_3$ to the linear effect of the stirrer rotation and $ß_4$ to the linear effect of the pump power. The terms $ß_{11}$ to $ß_{22}$ refer to the respective quadratic effects, whereas the terms with mixed indices (e.g. $ß_{12}$) refer to the respective interactions of the process variables towards the target values. For each model derivation method, the Tables 5.10, 5.12 and 5.14 exhibt, which model terms (effect, interaction) have been identified as significant by notation of the corresponding parameter value ß. Terms that have not been denoted with a parameter value ß, have not been identified as being significant by statistical means.

Model	Method of significance testing	$ß_0$	$ß_1$	$ß_2$	$ß_3$	$ß_4$	$ß_{11}$	$ß_{22}$	$ß_{33}$	$ß_{44}$	$ß_{12}$	$ß_{13}$	$ß_{14}$	$ß_{23}$	$ß_{24}$	$ß_{34}$	Validation
1st grade	1 / a	9,05		-2,77		1,74									1,00		No
	1 / a	9,05		-2,77		1,74									1,00		No
	1 / b	9,05		-2,77		1,74				-0,67					1,00		No
2nd grade	2 / a	9,05		-2,77		1,74	0,53		0,50	-0,67					1,00		No
	2 / b	9,05	0,22	-2,77		1,74	0,53	-0,39	0,50	-0,67					1,00		No
	3 / a	9,05		-2,77		1,74									1,00		No
	3 / b	9,05		-2,77		1,74	0,53		0,50	-0,67					1,00		No

Table 5.10: Derivation of 1st and 2nd grade models by RSM for rancimat value

Below, the applied methods for significance testing and validation are specified.

- Method 1/a : Estimation of error variance from negligible model terms and determination of parameter significance by relative variance (F-Test) .

- Method 1/b : Estimation of error variance from negligible model terms and determination of parameter significance by individual confidence intervals (t-test).

- Method 2/a : Estimation of error variance from multiple experiment realisation and determination of parameter significance by relative variance (F-Test) .

- Method 2/b : Estimation of error variance from multiple experiment realisation and determination of parameter significance by individual confidence intervals (t-test).

- Method 3/a : Estimation of error variance from the assumption that the model is valid and determination of parameter significance by relative variance (F-Test).

- Method 3/b : Estimation of error variance from the assumption that the model is valid and determination of parameter significance by individual confidence intervals (t-test).

- Method model validation RSM : F-Test on ratio of summed squares.

The objective of applying the RSM is just to identify statistically significant model terms, process models respectively, based on the results of the SDoE. The 1st grade model and the

2nd grade model determined by method 2/b are then parameterized (fit to the experimental data, respectively), validated and analyzed by Bayesian means in 5.2. The 2nd grade model has been chosen exemplarily for further analysis as the analysis and comparison of all derived models is beyond the scope of this work.

5.1.2.4 Model derivation for target value acid value

Table 5.11 shows the results for two repetitions of the orthogonal CCD for the rancimat value. The model derivation has been achieved on the mean values and without definition and exclusion of outliers. This procedure should provide evidence that the RSM and Bayesian approach applied here is capable of handling even experimental data that is significantly interfered by noise.

Experiment		1	2	3	4	5	6	7	8	9	10	11	12	13
Acid value	Run 1	0,33	0,36	0,33	0,28	1,6	0,57	1,56	0,87	0,29	0,25	0,2	0,32	0,69
$[mg\,KOH{*}kg^{-1}]$	Run 2	0,27	0,26	0,21	0,26	0,78	0,55	1,57	0,45	0,28	0,19	0,24	0,32	0,54
Experiment		14	15	16	17	18	19	20	21	22	23	24	25	
Acid value	Run 1	0,48	1,02	0,28	0,47	0,26	0,38	1,39	0,49	0,35	0,44	0,45	0,34	
$[mg\,KOH{*}kg^{-1}]$	Run 2	0,23	0,63	0,31	0,38	0,23	0,19	0,81	0,31	0,33	0,41	0,3	0,32	

Table 5.11: Results orthogonal central composite design for acid value

Table 5.12 exhibits the derivation of 1st and 2nd grade process models with significant model terms, which have been determined by six different methods of statistical significance testing. The model derivation and validation has been achieved according to the methods exposed theoretically in chapter 2.3.2.1. Also compare 5.1.2.3 for the significance testing and validation methods applied here. It can be observed that these methods yield three different process models with statistically significant model terms. All models are valid according to the statistical validation method applied here.

Model	Method of significance testing	$ß_0$	$ß_1$	$ß_2$	$ß_3$	$ß_4$	$ß_{11}$	$ß_{22}$	$ß_{33}$	$ß_{44}$	$ß_{12}$	$ß_{13}$	$ß_{14}$	$ß_{23}$	$ß_{24}$	$ß_{34}$	Validation
1st grade	1 / a	0,52	0,12	-0,24		0,14					-0,11				-0,15		Yes
	1 / a	0,49	0,11	-0,25		0,12		0,16			-0,11				-0,15		Yes
	1 / b	0,49	0,11	-0,25		0,12		0,16			-0,11		0,05	0,04	-0,15		Yes
2nd grade	2 / a	0,49	0,11	-0,25		0,12		0,16			-0,11				-0,15		Yes
	2 / b	0,49	0,11	-0,25		0,12		0,16			-0,11				-0,15		Yes
	3 / a	0,49	0,11	-0,25		0,12		0,16			-0,11				-0,15		Yes
	3 / b	0,49	0,11	-0,25		0,12		0,16			-0,11				-0,15		Yes

Table 5.12: Derivation of 1st and 2nd grade models by RSM for acid value

The objective of applying the RSM is just to identify statistically significant model terms, process models respectively, based on the results of the SDoE. The 1st grade model and the 2nd grade model determined by method 1/a are then parameterized (fit), validated and analyzed by Bayesian means in chapter 5.2. The 2nd grade model has been chosen exemplarily for further analysis as the analysis and comparison of all derived models is beyond the scope of this work.

5.1.2.5 Model derivation for target value tocopherol

Table 5.13 shows the results for two repetitions of the orthogonal CCD for the tocopherol content. The model derivation has been achieved on the mean values and without definition and exclusion of outliers. This procedure should provide evidence that the RSM and Bayesian approach applied here is capable of handling even experimental data that is significantly interfered by noise.

Experiment		1	2	3	4	5	6	7	8	9	10	11	12	13
Tocopherol	Run 1	10,52	1,25	15,46	1,10	43,99	32,23	45,94	34,99	9,26	1,01	17,08	29,39	45,38
$[mg*100g^{-1}]$	Run 2	11,32	0,98	15,87	1,06	44,03	32,22	47,62	37,25	8,63	0,91	15,04	0,98	43,00
Experiment		14	15	16	17	18	19	20	21	22	23	24	25	
Tocopherol	Run 1	28,92	42,98	29,39	23,06	16,88	0,94	42,83	17,35	23,30	31,58	2,30	24,70	
$[mg*100g^{-1}]$	Run 2	29,30	45,13	32,64	24,66	20,39	1,04	45,33	16,87	25,72	36,62	1,35	17,87	

Table 5.13: Results orthogonal central composite design for tocopherol

Table 5.14 exhibits the derivation of 1st and 2nd grade process models with significant model terms, which have been determined by six different methods of statistical significance testing. The model derivation and validation has been achieved according to the methods exposed theoretically in chapter 2.3.2.1. Also compare 5.1.2.3 for the significance testing and validation methods applied here. It can be observed that these methods yield three different process models with statistically significant model terms. Except the 2nd grade model determined by method 3/a, all models are valid according to the statistical validation method applied here.

Model	Method of significance testing	Model term															Validation
		β_0	β_1	β_2	β_3	β_4	β_{11}	β_{22}	β_{33}	β_{44}	β_{12}	β_{13}	β_{14}	β_{23}	β_{24}	β_{34}	
1st grade	1 / a	23,59		-14,85	-2,16	5,24											Yes
	1 / a	22,55		-14,93	-2,25	6,47		2,07			-1,45						Yes
	1 / b	22,55		-14,93	-2,25	6,47	1,42	2,07			-1,45			-1,1	-1,1		Yes
2nd grade	2 / a	23,59		-14,85	-2,16	5,24											Yes
	2 / b	23,59		-14,85	-2,16	5,24											Yes
	3 / a	23,59		-14,85		5,24											No
	3 / b	23,59		-14,85	-2,16	5,24											Yes

Table 5.14: Derivation of 1st and 2nd grade models by RSM for tocopherol

The objective of applying the RSM is just to identify statistically significant model terms, process models respectively, based on the results of the SDoE. The 1st grade model and the 2nd grade model determined by method 1/a are then parameterized (fit), validated and analyzed by Bayesian means in chapter 5.2. The 2nd grade model has been chosen exemplarily for further analysis as the analysis and comparison of all derived models goes beyond the scope of this work.

5.1.2.6 General observations during the trials

After the short path distillation (SPD) process, deodorized palm oil exhibited almost the same dark orange color as raw palm oil. For comparison see samples 1 and 3 (left to right) in Figure 5.4. The color is an indicator, that no separation of carotene from the crude palm oil took place. In contrast, during the standard deodorization, the carotene is heat induced decomposed and its volatile decomposition products are removed from the palm oil. The standard deodorized palm oil can be seen in sample 4 of Figure 5.4 on the right site (Rimbach et al. 2010, p. 181; Bockisch, 1993, pp. 557–558).

Figure 5.4: Color comparison of palm oil after various refining stages

During the experiments, the dependency of the output of palm oil distillate on the feed flow rate was analyzed. The amount of distillate increased with decreasing feed flow rate. For the evaporator temperature the correlation with the process output was even more significant. With an increasing evaporator temperature also the volume of distillate remarkably increased.

With this experimental setup, the color of the distillate became darker orange with each additional distillation cycle. This indicates an increasing separation of carotene from the oil (Rimbach et al., 2010, p. 174). In the first cycle of the distillation, a bright yellow distillate was deposited on the condenser. This sample then became darker with each distillation cycle. In the final cycle, the distillate reached an orange color that was comparable to the distillation residue. Through further analysis of the residues could be confirmed, that it still contained some degree of carotene. Finally, the beginning of diacylglycerine (DAG) separation was confirmed analytically for the experiments of the SDoE that were achieved at the highest evaporator temperatures.

In contrast, with a minimum evaporator temperature and a maximum feed flow rate (according SDoE), the output of distillate decreased from 17 g to 13 g. At evaporator temperatures of below 150°C the distillate became bright yellow (Figure 5.4, second from right).

As expected, the acid value of the distillation residue was reduced with each additional distillation cycle. The most intensive decrease of the acid value was found after the first cycle, with a lower feed flow rate leading to a greater decrease. At a condenser temperature of below 50°C, the condensate was dispersed in solid form (see Annex A – SPD).

Also in the cold trap prior to the vacuum pumps, very small quantities (0.1 g) of a yellow distillate with a very intensive and pungent odor could be observed. In the literature is assumed that these are volatile flavors and fragrances (Bockisch, 1993, pp. 74–78). Figure 5.5 shows a color comparison between a bright distillate on the left and a residue on the right gained at an evaporation temperature of 150°C

Figure 5.5: Comparison of distillation residues and condensates

The pressure in the empty and evacuated plant was 10^{-3} mbar, which eventually increased to 10^{-2} mbar when the palm oil was first introduced into the evacuated plant in the first cycle. In the following distillation cycles, the pressure decreased again to the starting value of 10^{-3} mbar (see also Annex a – pictures of an SPD plant).

5.2 Bayesian model fit, discrimination and performance analysis

A 1st grade model is based on a smaller full factorial design that enables the identification of the linear effects of process variables towards the target values, as well as the identification of interactions (up to 3rd order in case of four process variables). A Bayesian model parameterization and performance analysis of a 1st grade model, which is based only on the experimental data of the full factorial design, also provides results that just refer specifically to that particular SDoE and data. For example, a valid model is only valid for the estimation and simulation of the linear effects and interactions that can be determined from this particular SDoE.

One primary reason for applying Bayesian model fit, discrimination and analysis methods is to clarify the question of whether the improved performance of a 2nd grade model justifies a higher experimental and computational effort in comparison to a 1st grade model. To answer this question, a 1st grade model must therefore also be applied to the data set of the orthogonal CCD with respect to Bayesian model fit, validation and discrimination. In order to correctly calculate the relative model probability, both alternative models have to be applied to the same data set. Otherwise, the first grade model would naturally be favored due to a higher likelihood of the candidate vector which is caused by a much smaller number of data points/experimental settings involved in the likelihood calculation (Eq. 2.23).

If a 1st grade model is parameterized (fit) and valid with respect to the orthogonal CCD data, it can be stated, that this model is sufficient for simulating the true underlying process. In other words, the consideration of existing quadratic effects through additional experimental and computational effort by an orthogonal CCD is not crucial for simulating the true underlying process with sufficient accuracy. This is true because these quadratic effects are negligible compared to the linear effects. The latter have already been identified by the smaller full factorial design. Such a situation could theoretically occur along with a relative model probability that clearly favors a more complex 2nd grade model.

It is a well-accepted fact, that in most cases 2nd grade models are more flexible due to a larger number of model terms, which yields a better data fit by nature. As previously mentioned, slightly increased likelihoods for single data points or experimental settings can cause substantial differences between the results of 1st and 2nd grade models regarding the likelihood of the total data set (the overall likelihood of a data set is the product of the individual likelihoods for the single experimental settings). This, in turn, might significantly

influence the relative model probability. However, this does not necessarily mean that the 1st grade model is insufficient.

In this work, all prior distributions, likelihoods and proposal densities applied are assumed to be Gaussian. According to the theory and principles applied (compare chapter 2.3.2.2), 1000 Markov-Chains were run in parallel with 20 update cycles of the proposal density each. Within each update cycle 10 tuning cycles of the proposal density were achieved with 20 chain steps within each tuning cycle (4000 chain steps in total for each Markov-Chain). The Bayesian law (Eq. 2.18) shows, that the posterior probability is affected by the prior probability. The more distinct the prior knowledge (low standard deviation) is, the higher is its influence on the posterior probability distribution.

5.2.1 Implementation and validation of the Bayesian methodology by the example of a growth model

The objective of implementing the Bayesian methodology at first for the simple example of a microbial growth model is to verify the correct understanding and implementation of the mathematical theory, as well as the analysis of the method dynamics and results. This procedure provides a reliable basis for an efficient method application and adjustment to the real life problem of model based SPD optimization. It also enables a correct interpretation of the method results.

It has to be initially mentioned at this point, that all following simulation and model analysis experiments have been achieved by application of the posterior mean (estimate of expected value) and mode as Bayesian point estimates for model parameterization and candidate vector for model discrimination. Actually, these alternatives only affect the result of the model discrimination. The model validation and distribution of experimental observations, are not affected by that choice. Thus, only for the relative model probability alternative results occur and are presented in the tables below. The relative model probability for the mode as candidate vector is presented in brackets.

5.2.1.1 Does the implemented method provide chain convergence and true underlying process recovery?

Does the implemented theory produce a chain convergence from a prior to a posterior distribution?

How far is the method / algorithm able to recover the true model, parameterization and observation from a direct solution of this known model (non-interfered data) when starting the MCMC-process from a parameterization much different from the real one?

To answer these questions a simulation experiment was achieved in which the Bayesian methodology for model fit, discrimination and analysis was applied to a simulated experimental data set, which was derived as direct solution from the so-called Moser-model with a defined known parameterization. This experiment is denoted as 'Solution Moser 2'. Table 5.15 shows the conditions for this simulated experimental data generation. (The model parameter values applied here are the results of an RLS-fit to the actual experimental data in Table 5.24 which are rounded up to the second decimal place in Tabel 5.15. In practice, all further computations were achieved with the non-rounded data output of the program / algorithms.)

Generation of model solution				
Model	Parameters for model solution			
	μ_{max}	K_S	Y_{XS}	n
Moser	0,46	21,50	0,48	2,50
Michaelis-Menten	0,47	14,17	0,48	

Random addition of noise to model solution		
Processvariable	Standard deviation for sampling *[%]*	
	Direct solution	Random sampling
Cell concentration	0	0,1
Substrate concentration	0	0,05

Table 5.15: Initial conditions of simulated experimental data generation

Table 5.16 shows the corresponding simulated experimental data to which the Bayesian methodology of model fit, discrimination and analysis was applied. This data is the basis of the experiment 'Solution Moser 1' as well, which is exposed in the next sub-chapter. For presentation purposes, the data in Table 5.16 was rounded up to the second decimal place. All computations were achieved with the non-rounded data output from the random sampling of the simulated experimental data.

Data source	Process variable	Data											
	Time *[h]*	0	1	2	3	4	4,5	5	5,5	6	6,5	7	8
Solution Moser 1 und 2	Cell conc. $[g*L^{-1}]$	10,00	15,79	24,92	39,33	62,11	78,01	97,86	121,62	131,86	130,82	130,82	130,82
	Substrate conc. $[g*L^{-1}]$	200,00	188,02	169,10	139,27	92,09	59,17	18,06	1,00	1,00	1,00	1,00	1,00

Table 5.16: Simulated experimental data: Direct solution from Moser-Model

Table 5.17 exhibits the initial conditions and results of the Bayesian model fit and discrimination with respect to the simulated experimental data in Tabel 5.16. Initial conditions are the mode and variance of the prior parameter distribution, which are applied as prior knowledge about the parameterization of the Michaelis-Menten- and Moser-model.

This distribution is assumed to be Gaussian, where distribution mode and expected value are equal. Results are the parameterization and residual least squares (RLS) of a RLS-fit, as well as the Bayesian point estimates, posterior distribution parameter estimates respectively, which are the estimates of distribution variance, mode and expected value . Additionally the Bayesian relative model probability is exposed (last column). The values exposed in Table 5.17 are rounded -up to the second decimal place for space reasons. All computations were achieved with non-rounded data output of the programs / algorithms.

Data source	Model	Prior knowledge[*]					Method	Parameterization					Discrimination	
								Specific value	Parameter					
			μ_{max}	K_S	Y_{XS}	n			μ_{max}	K_S	Y_{XS}	n	RLS	P_{MM}/P_{Moser}
	Michealis-	Parameter	μ_{max}	K_S	Y_{XS}		Bayes	Parameter**	0,47 (0,46)	2,31 (0,64)	0,50 (0,48)			
	Menten	Mode	0,46	0,60	0,48			Variance	$3,23*10^{-4}$	5,81	$7,20*10^{-4}$		0,24	
Solution		Standard dev. *[%]*	25	25	25		RLS	Parameter	0,46	0,60	0,48			0,09 (0,1)
Moser 2		Parameter	μ_{max}	K_S	Y_{XS}	n	Bayes	Parameter**	0,46 (0,46)	21,05 (22,77)	0,49 (0,49)	3,05 (2,59)		
	Moser	Mode	0,46	21,50	0,48	2,50		Variance	$7,51*10^{-4}$	118,26	$4,98*10^{-4}$	0,89	$1,68*10^{-14}$	
		Standard dev. *[%]*	25	25	25	25	RLS	Parameter	0,46	21,50	0,48	2,50		
* arbitrarily chosen for analytic reasons; dimension of mode is based on RLS fit to actual exp. Data														
**no brackets: posterior expected value estimate (mean); in brackets: posterior mode estimate														

Table 5.17: Initial conditions and results of model fit and discrimination towards a direct solution from a Moser-Model – Approach 2

Table 5.18 shows the parameter start values of the Markov-Chains. From that start values the Markov-Chains converge towards the posterior probability distribution with respect to the prior knowledge on model parameterization and the likelihood.

Model	Start values of Markov-Chains				
Michealis-Menten	Parameter	μ_{max}	K_S	Y_{XS}	
	Start value	0,6	18	0,6	
Moser	Parameter	μ_{max}	K_S	Y_{XS}	n
	Start value	0,6	24	0,6	1,5

Table 5.18: Start values of Markov-Chains

Answering the two initial questions of 5.2.1.1 is important in order to prove, that the mathematical theory is understood and correctly implemented by a functional / precise algorithm. For this purpose the prior knowledge has to be as neutral as possible. If it is, in contrast, exactly defined and much different from the true parameterization (which is known in this constructed testing scenario) it would drive the final solution away from the true one, which shall be recovered here from the data only and without interference from the assumed prior knowledge. Thus, the chosen prior is a distribution, which is centered on the RLS-fit parameterization (equal to the parameterization of the true underlying model from which the simulated experimental data was derived) with a comparatively large standard deviation of 25% (compare Table 5.15 and 5.17). If the Bayesian method applied

here is implemented and functioning correctly, then the information about the true underlying parameterization is extracted from the data and the Markov-Chains converge towards the known true parameterization with decreasing variance for sensitive parameters (compare explanation of this effect below). In the opposite case a significantly different posterior probability distribution (especially with respect to the distribution mode and expected value) would emerge. As true underlying model the Moser-Model was applied here with the data and conditions exposed in Table 5.15 and 5.16. Table 5.17 shows the results of the Bayesian model fit and discrimination. It has to be underlined, that the mean (estimate of distribution expected value) of the posterior distribution sample (values without brackets) differs from the mode estimate (values in brackets) for both models, especially for the parameters K_s and n (Moser-model). This inidicates, that even though a Gaussian prior and likelihood is applied (mode and expected value are equal in this case), the posterior distribution might exhibit an asymmetric non-Gaussian shape. It is important to mention in that context that Chib and Jeliazkov recommend the parameterization with the highest posterior probability as candidate vector for model discrimination, as this might significantly influence the accuracy of relative model probability estimation (Chib, S. and Jeliazkov, I., 2001, pp. 270-281). Thus, the distribution mode and expected value (their estimates respectively) are differentiated explicitly in this work. The results in Table 5.17 show, that the parameter values for μ_{max} and Y_{XS} of the true underlying Moser-Model are recovered precisely by the estimates of the posterior distribution mode and expected value, whereas the posterior mode estimate for K_S differs by 6% and the expected value estimate of n by 22% from the true values. It also has to be underlined here, that for μ_{max} and Y_{XS} the posterior variance is very small (high certainty of prediction), whereas it is very large for K_S and n (high uncertainty of prediction). This is not an indication for an insufficient implementation / functioning of the method. The variance of the Bayesian posterior distribution is a fundamental output of the Bayesian method implemented here and defines the range of parameter values with significant possibility in the light of the data. This range does not only depend on the error of experimentation applied, but also on the parameter sensitivity. This means that for model parameters with comparatively low sensitivity (comparatively low influence on model solutions), also a comparatively broad range of parameter values produces similar model solutions and therefore also exhibits a comparatively high probability of approaching the true value. Thus, the posterior variance is also an indicator for parameter sensitivity.

The Bayesian method applied here produces variances of the posterior distribution by nature, as it assumes a defined error of experimentation for the experimental data (even if it was drawn directly from a defined known model as done in this example). Thus, the results for the parameters K_S and n in Table 5.17 are quite plausible and do not disagree

with correct method implementation and method performance. On the contrary, these results underline the capability of the method to provide proper uncertainty measures, which result from the explicit consideration of errors of experimentation that might interfer the experimental data.

Figures C.9 to C.12 in Annex C examplarly show the convergence of the distribution mean and variance of the Michaelis-Menten-Model (MM-Model) and Moser-Model parameters. A steady convergence of the Markov-Chains from the start values to the posterior means can be clearly observed with a stable posterior mean (point of completed convergence) at least from the 10th (a=10 in program) update of the proposal density (estimation of the covariance matrix, respectively). With a relative model probability of P_{MM}/P_{Moser}=0,09985 (posterior mode estimate as candidate vector) the Moser-Model could be identified as being significantly more probable in the light of the data and thus could be recovered as true underlying model. The application of the posterior mean (expected value estimate respectively) as candidate vector for model discrimination produces similar results in this case. The conventional approach of comparing RLS for model choice exhibits a much more distinct decision for the Moser-Model. This is caused by the fact that by application of RLS, the model is directly fit to the crude experimental data (which is a direct model solution in this case), whereas the Bayesian approach considers the potential interference of that data by experimentation noise. The latter approach thus also assigns a certain probability to alternative data sets that in turn assign certain probabilities to the competing Michaelis-Menten-Model (MM-Model). By nature, this process yields a less distinct favouring of the Moser-Model by Bayesian relative model probability estimation, which is in practice more realistic due to explicit consideration of experimentation noise and, thus, alternative events. Due to the observations exposed so far it can be stated, that the Bayesian method applied here produces converging Markov-Chains and is capable of identifying the true underlying model and parameterization from non-interferred data at completely different starting values.

Figure C.13 to C.30 in Annex C show for both models the prior and posterior parameter distribution (exemplary for one parameter due to space reasons; similar for other parameters) and model fit, as well as the model validation and posterior probability of experimental observations according to the method proposed by Geweke (Geweke, 2007, pp. 3529-3550). According to Figure C.19 to C.26 both models are accepted as being valid as in all cases the experimental value is in support of the distribution (within the 5%- and 95%-quantile). This means, that both models approach the true underlying process with the desired accuracy.

Figures C.27 to C.30 examplaryly show the posterior probability of experimental observations for the timepoint t=5 h for the Michaelis-Menten- and Moser-Model. For both models, the experimental value (which is a direct solution of the Moser-Model assumed as true underlying process) is located in the centre of the distribution and is therefore in excellent support of the distribution. Thus, it can be stated that both models proofed to be suitable for a precise estimation of the experimental observations to be expected. For the Moser-Model the experimental value matches the posterior mean whereas it differs slightly for the MM-Model. This observation / result underlines the capability of the Bayesian method applied here to estimate the true event with sufficient precision and to identify the most probable true underlying process. Additionally, it provides a very usefull measure of prediction uncertainty by means of posterior distributions. The uncertainty / distribution is subject to the error of experimentation applied to the Bayesian process. It has to be mentioned here, that a broad range of other experimental observations could possibly occur if the applied model was true due to the fact that also a broad range of other experimental data sets could possibly occur due to the assumed error of experimentation.

Due to the above mentioned results, analysis and interpretation, it can be stated that the Bayesian method theoretically exposed in chapter 2.3.2.2 has been implemented correctly by algorithms and exhibits the requested performance in terms of the analytic measures / output and the predictions that can be derived from that.

5.2.1.2 What is the influence of prior knowledge?

In which way does the choice of the prior knowledge influences the analytic results and how can incorrect / inadequate prior knowledge superimpose the information about the true underlying process contained in the experimental data?

To answer this questions, two simulation experiments were carried out in which the Bayesian methodology for model fit, discrimination and analysis was applied to the same simulated experimental data set that was derived as direct solution from the Moser-Model with a known parameterization (same data set as in 5.2.1.1, see Tabel 5.16). The two simulation approaches, which are denoted as 'Solution Moser 1' and 'Solution Moser 2' (the latter is also subject to the preceding sub-chapter), only differ in the choice of the prior mode (equal to expected value as prior distribution is assumed to be Gaussian) as prior knowledge on the model parameterization. Table 5.15 shows the conditions for this simulated experimental data generation and Table 5.16 the corresponding simulated experimental data. The start values of the Markov-Chains were equal for all simulation experiments and are exposed in Table 5.18. Table 5.19 exhibits the initial conditions and results of two different approaches of Bayesian model fit and discrimination towards the

direct solution from a Moser-Model (Tabel 5.15 and 5.16). The approaches only differ in the choice of the prior mean.

Data source	Model	Prior knowledge*					Method	Parameterization						Discrimination	
								Specific value	Parameter						
			μ_{max}	K_S	Y_{XS}	n			μ_{max}	K_S	Y_{XS}	n		RLS	P_{MM}/P_{Moser}
	Michealis-	Parameter	μ_{max}	K_S	Y_{XS}		Bayes	Parameter**	0,50 (0,49)	6,12 (5,61)	0,53 (0,52)				
	Menten	Mode	0,6	18	0,6			Variance	$4,06*10^{-4}$	9,3	$7,64*10^{-4}$			0,24	
Solution		Standard dev. *[%]*	25	25	25		RLS	Parameter	0,46	0,6	0,48				0,01 (0,01)
Moser 1		Parameter	μ_{max}	K_S	Y_{XS}	n	Bayes	Parameter**	0,47 (0,46)	22,39 (22,81)	0,51 (0,49)	1,84 (1,85)			
	Moser	Mode	0,6	24	0,6	1,5		Variance	$7,83*10^{-4}$	40,065	$6,12*10^{-4}$	0,079		$1,68*10^{-14}$	
		Standard dev. *[%]*	25	25	25	25	RLS	Parameter	0,46	21,5	0,48	2,5			
	Michealis-	Parameter	μ_{max}	K_S	Y_{XS}		Bayes	Parameter**	0,47 (0,46)	2,31 (0,64)	0,50 (0,48)				
	Menten	Mode	0,46	0,60	0,48			Variance	$3,23*10^{-4}$	5,81	$7,20*10^{-4}$			0,24	
Solution		Standard dev. *[%]*	25	25	25		RLS	Parameter	0,46	0,60	0,48				0,09 (0,1)
Moser 2		Parameter	μ_{max}	K_S	Y_{XS}	n	Bayes	Parameter**	0,46 (0,46)	21,05 (22,77)	0,49 (0,49)	3,05 (2,59)			
	Moser	Mode	0,46	21,50	0,48	2,50		Variance	$7,51*10^{-4}$	118,26	$4,98*10^{-4}$	0,89		$1,68*10^{-14}$	
		Standard dev. *[%]*	25	25	25	25	RLS	Parameter	0,46	21,50	0,48	2,50			

*arbitrarily chosen for analytic reasons; dimension of mode is based on RLS fit to actual exp. Data (Sol. Moser 2) or MC-start value (Sol. Moser 1)
**no brackets: posterior expected value estimate (mean); in brackets: posterior mode estimate

Table 5.19: Initial conditions and results of model fit and discrimination towards a direct solution from a Moser-Model – Approach 1 and 2

The results of the second approach, which is denoted as 'Solution Moser 2', are exposed in Annex C and have been already discussed in 5.2.1.1. For the first approach, which is denoted as 'Solution from Moser 1', the results for Markov-Chain convergence, posterior parameter distribution, comparison of model fits, model validation and posterior distribution of observations are exposed in Annex D.

The analysis achieved here is crucial for the right choice of the prior knowledge in the analysis of the SPD process. So how does the prior knowledge applied to the Bayesian methodology influence the analytic results? The Bayesian process implemented here starts with a defined parameter start vector which is identical for both approaches. From a proposal density, which is centered on this start vector, a so called candidate vector is drawn by random sampling. The relative posterior probability of the current state (equals start vector in the first process step) and the proposal is then calculated according Eq. 2.21 and 2.20 from the prior probabilities and the likelihoods. Dependent on the relative posterior probability, proposals are accepted as new states of the Markov-Chain according to the procedure exposed in chapter 2.3.2.2.2. At chain convergence all accepted states form a representative sample of the posterior parameter distribution. This means, that for a defined parameter vector, in comparison to other vectors, the probability to be accepted as

part of the posterior sample (Eq. 2.21), and thus also the posterior probability distribution, depends by equal parts on the ratio of prior probabilities and likelihoods. Equation 2.21 shows, that in case of equal prior probabilities (this equals no prior knowledge on model parameterization) the prior probability ratio can be eliminated which reduces the posterior probability ratio to the likelihood ratio. In this case the posterior distribution is only subject to the information contained in the experimental data. In contrast to that, this information can be superimposed by contrary prior knowledge, if a likelihood ratio that favors a certain parameter vector faces a prior probability ratio that favors the other vector. This shows that the prior knowledge to be applied has to be wisely chosen by terms of prior mode (equals prior expected value in case of Gaussian distribution) and uncertainty of this knowledge (prior variance). A careless choice and application of prior knowledge can nullify the entire experimental information and effort. The larger the variance of the prior distribution, the more the parameter vectors equal in prior probability and the less prior knowledge is applied to the Bayesian process. The importance of the prior mode positioning therefore declines with increasing variance. Contrary to that, the difference in prior probability between parameter vectors increases with decreasing variance. In this case, also the importance of prior mode positioning, as well as of the prior knowledge applied, increases. To illustrate the effect of prior probability choice and application, the experiment denoted 'Solution Moser 1' is achieved and compared to the experiment 'Solution Moser 2', which just differs in the choice of the prior mode (compare Tabel 5.19). For the experiment 'Solution Moser 1' the prior mode is shifted to the start values of the Markov-Chains, which differ significantly from the known real parameterization (compare Tabel 5.15). The data in Table 5.19 shows for both models and for all parameters, that by shifting the mode of the prior knowledge in a defined direction, also the posterior mode and expected value estimate is shifted in this direction. By choosing the above mentioned start values as prior mode ('Solution Moser 1'), this shift has been achieved much more crucial for the Michaelis-Menten-Model. This means, that for this model the information on model parameterization contained in the data is much more superimposed by the imprecise prior knowledge applied. Indeed, for the Michaelis-Menten-Model the applied prior mode for parameter K_S differs much more from the parameterization to be expected (e.g. RLS-fit parameterization), compared to the Moser-Model. The comparison of the relative model probabilities also shows, that additionally this asymmetric application of bad prior knowledge to the different models also effects model discrimination. A model upon which more impreciseness in prior knowledge is imposed loses probability. In the study presented here, the relative model probability is changed by the power of ten due to the application of more imprecise prior knowledge to the Michaelis-Menten-Model in experiment 'Solution Moser 1'. Figure D.35 and D.36 show that the Bayesian fit of the

Michaelis-Menten-Model to the experimental data is less precise compared to the Moser-Model and the experiment 'Solution Moser 2' (especially for the cell concentration). This is caused by the above mentioned assymetric superimposition of the information contained in the experimental data by imprecise prior knowledge, which discriminates the Michaelis-Menten-Model. This causes the situation that the experimental value for the cell concentration is not in support of the distribution of experimental observations that could occur if the Michaelis-Menten-Model was true (Figure D.49). Nevertheless, both models are validated by the experimental values for solution mean and variance that are in very good support of the distributions (Figure D.41 to D.48). Thus by Bayesian means, despite of favouring the Moser-Model by more precise prior knowledge, both models describe the true underlying process with the desired accuracy. Figure D.31 to D.34 show an excellent stable convergence of the Markov-Chains from the start vectors to the posterior mean (converged from update 10 on). This indicates, that the parameter sample taken from the Markov-Chains is a valid representative of the posterior distribution. Figure D.37 to D.40 exemplarly show for both models and for a defined single parameter the efficiency of the Bayesian process of updating the distribution mean (estimate of expected value) and narrowing the prediction uncertainty (variance) by extracting the information on the true underlying process contained in the experimental data.

5.2.1.3 How to interpret the output of the Gibbs-sampling?

How has the output of the Gibbs-sampling, the Markov-Chains respectively, to be interpreted in terms of determination of the posterior parameter distribution and model parameterization (model fit respectively)?

For this analysis the simulation experiment denoted as 'Random sampling Michaelis-Menten' was achieved, in which simulated experimental data was generated by adding random noise to a direct solution of the Michaelis-Menten-Model. Table 5.15 shows the conditions of simulated experimental data generation and Table 5.18 the start vectors of the Markov-Chains, which were equal for all simulation experiments. In Table 5.20 the resulting simulated experimental data is exposed. For presentation reasons it has been rounded-up to the second decimal place. All further computations were achieved with the non-rounded data output of the program / algorithm.

Data source	Process variable	Data											
Random sampling Michaelis-Menten	Time *[h]*	0	1	2	3	4	4,5	5	5,5	6	6,5	7	8
	Cell conc. $[g*L^{-1}]$	9,64	16,80	28,06	35,89	51,50	59,93	81,04	106,12	113,01	99,99	114,63	102,23
	Substrate conc. $[g*L^{-1}]$	190,01	179,89	169,20	138,63	105,73	81,12	51,33	20,60	1,04	1,10	0,98	0,95

Table 5.20: Simulated experimental data: Random sampling from Michaelis-Menten-Model

Table 5.21 exhibits the initial conditions and results of model fit and discrimination with respect to this simulated experimental data in Tabel 5.20.

Data source	Model	Prior knowledge[*]					Method	Parameterization					Discrimination	
								Specific value	Parameter					P_{MM}/ P_{Moser}
									μ_{max}	K_S	Y_{XS}	n	RLS	
	Michaelis-Menten	Parameter	μ_{max}	K_S	Y_{XS}		Bayes	Parameter[**]	0,45	5,75	0,53			
Random		Mode	0,48	15,31	0,51			Variance	$2,12*10^{-4}$	8,18	$4,85*10^{-4}$		481	
sampling		Standard dev. [%]	50	50	50		RLS	Parameter	0,48	15,31	0,51			23,36
Michaelis-	Moser	Parameter	μ_{max}	K_S	Y_{XS}	n	Bayes	Parameter[**]	0,52	3,24	0,51	0,53		
Menten		Mode	0,44	5,85	0,54	2,18		Variance	$1,7*10^{-3}$	8,52	$8,83*10^{-4}$	1,32	501,7	
		Standard dev. [%]	50	50	50	50	RLS	Parameter	0,44	5,85	0,54	2,18		
* mode is based on RLS fit to simulated exp. Data														
** mode estimate of posterior distribution														

Table 5.21: Initial conditions and results of model fit and discrimination towards a random sample from the Michaelis-Menten-Model

All figures regarding this subchapter and experiment are comprised in Annex E.

For the simulation experiment denoted as 'Random sampling Michaelis-Menten' the calculation of the parameter posterior mean (estimate of posterior expected value) from the Markov-Chain output for the Moser-Model (at chain convergence) results in a parameter vector that exhibits a comparatively very low likelihood (probability that the exp. data occurs given the Moser-Model with that parameterization (with respect to error in experimentation)), which is significantly lower than the likelihoods of most parameter vectors in the posterior sample, especially in comparison to the mode estimate. Studies showed, that for this simulated experimental data set especially slight changes in the parameter n of the Moser-Model cause significant changes of the likelihood, which can be referred to regions of the experimental data that are comparatively highly affected by errors in experimentation (in the example presented here, these are the last five data points of the cell concentration). Thus, inaccuracies in estimation of the parameter n significantly affect the likelihood of a parameter vector and corresponding modeling results. Actually, it could be shown, that the mean of parameter n calculated from the Markov-Chain output significantly differs from the value of the parameter vector with the highest posterior probability (posterior mode estimate). In fact, this is the reason for the comparatively very low likelihood observed for the parameter mean. The analysis of the data set 'Random sampling Michaelis-Menten' therefore reveals, that the mean and variance calculated from a representative distribution sample are suitable indicators for determination of Markov-Chain convergence (they are characteristic constants of a distribution that therefore stay stable at Markov-Chain convergence) but the mean is in some cases not identical with the posterior mode estimate. The latter is subject to the fact, that the mean of a representative posterior sample is just identical to the posterior mode, if

the posterior distribution was symmetric with respect to the sample mean (e.g. Gaussian distribution). Although the likelihood and prior distribution are assumed to be Gaussian in some cases, an asymmetric non-Gaussian posterior distribution might occur in which the sample mean might differ significantly from the sample mode (as in this case). For example, Table 5.19 above shows for some experiments significant differences between the mean and mode estimate (in brackets) of the posterior parameter distribution. Thus, the choice of distribution expected value and mode as point estimate for Bayesian model parameterization might exhibit significant differences in model prediction in case of assymetric posterior distributions. Chib and Jeliazkov also propose to apply parameter vectors with a high probability under the posterior distribution as candidate vectors for the estimation of the relative model probability, as it influences the precision of this estimation (Chib, S. and Jeliazkov, I., 2001, pp. 270-281). If for one model a candidate vector from a posterior region with high probability is chosen and for the other model a candidate vector from a low-probability region, then the former is automatically favoured in relative model probability estimation due to comparatively higher likelihoods. This case might occur, when the mean is chosen as candidate vector from an asymmetric posterior distribution. Indeed, it could be shown, that the choice of the posterior mean as candidate vector caused very extreme and unrealistic relative model probability values due to the extreme likelihood differences between the candidate vectors of Michaelis-Menten- and Moser-Model. The results in Table 5.19 also indicate, that the application of posterior mean and mode estimate as candidate vector can lead to different values for the relative model probability. These values do not differ in the general prediction regarding model choice but in the clearness of that prediction. An imprecise relative model probability estimate caused by suboptimal choice of the candidate vector might therefore also affect the result of Bayesian model averaging approaches in which simulations are subject to different models whose solutions are proportionally mixed with respect to the relative model probability. To allow for the above mentioned effects and results, the Bayesian point estimates for model parameterization, and thus the candidate vector for the relative model probability estimation, are derived from the posterior sample as mode estimate (most probable sample vector respectively) in further Bayesian analysis. For this purpose, for all parameter vectors of the posterior sample the product of likelihood and prior is calculated as relative measure for the posterior probability (the marginal likelihood is a normative constant being equal for all parameter vectors of the posterior sample; compare Eq. 2.18). For the above mentioned reasons, this principle of determination of Bayesian parameterization (candidate vector respectively) is applied to all further programming and analysis in this work as the most accurate and constructive approach. To provide further evidence for that approach, for some experiments in 5.2.2 also the application of the

distribution mean is achieved for analysis and interpretation reasons (as mentioned above). Figure E.53 to E.56 show for both models and all parameters the convergence of the distribution parameters mean and variance towards the posterior values. A stable convergence can be observed from the covariance update number 10 on. This shows again, that the Bayesian method applied here has been correctly implemented and produces a stable chain convergence. Thus, it is also a valid procedure to consider all parameter vectors from update 10 on as representative posterior sample. Figure E.57 to E.58 show the comparison of RLS- and Bayes-Fit for both models. It has to be emphasized that the RLS-fits do not show significant differences. Also the RLS-values which are a measure for conventional model discrimination approaches do not differ significantly (the Michaelis-Menten-Model is slightly favoured) whereas the Bayesian approach shows significant differences in model fit and clearly favours the Michaelis-Menten-Model with a relative model probability value of 23,36. This underlines that, in contrast to conventional model discrimination methods, the Bayesian methodology implemented here is capable of recovering the true underlying process from experimental data affected by. This is possible, as the Bayesian methodology explicitly considers errors of experimentation / noise by the sophicticated mathematical concept exposed in subchapter 2.3.2.2. Figure E.59 to E.62 exemplaryly show the prior and posterior parameter distribution for parameter 2 of the Michaelis-Menten-Model and parameter 3 of the Moser-Model. The validation of both models is given by Figure E.63 to E.70. Although the relative model probability clearly favours the Michaelis-Menten-Model both models are valid with respect to all Bayesian validation criteria applied here. This means, that both models are considered as being able of modeling the true underlying process with the desired accuracy. Figure E.71 to E.74 exemplaryly show the posterior probability of experimental observations at time point t=5 h for both models. For both models the experimental values are in good support of the distribution of values that are expected to occur if the models were true. This underlines the result of model validation, as both models predict observations that support the experimental values and are thus capable of modeling the true underlying process. Nevertheless, the experimental values are slightly better supported by the distributions of the Michaelis-Menten-Model which corresponds to the result of relative model probability estimation which favours the Michaelis-Menten-Model as being a more probable estimation of the true underlying process.

5.2.1.4 Is the method capable to recover the true process?

Is the method and algorithm implemented here able to recover the true underlying model and parameterization from a model solution that has been interfered by noise (standard deviation of experimentation) equivalent to that of the real experimental data?

To proof the suitability of method and algorithm in this sense, two different experiments have been achieved, in which experimental data has been simulated by adding random noise (equivalent to the error of experimentation of the real experimental data) to solutions of a defined Michaelis-Menten- and Moser-Model, respectively. The Bayesian method implemented here has been applied to this simulated experimental data in order to proof the ability to recover the underlying true model and parameterization which is known in this special case. Also the capability of the Bayesian method and algorithm to provide uncertainty measures of predictions has to be proofed and results to be interpreted.

Table 5.15 exhibits the conditions of direct model solution and simulated experimental data generation. Table 5.20 shows the simulated experimental data (random addition of noise) for the Michaelis-Menten-Model. Table 5.21 shows the initial conditions and results of Bayesian model fit and discrimination for the Michaelis-Menten- and Moser-Model with respect to that simulated experimental data. Annex E comprises all figures belonging to that experiment. Table 5.22 shows the simulated experimental data generated from the Moser-Model, which is subject to Bayesian model fit and discrimination.

Data source	Process variable	Data											
Random sampling Moser	Time $[h]$	0	1	2	3	4	4,5	5	5,5	6	6,5	7	8
	Cell conc. $[g*L^{-1}]$	9,21	15,16	23,40	35,66	63,22	74,66	99,97	144,29	127,19	133,22	142,14	144,39
	Substrate conc. $[g*L^{-1}]$	208,49	179,64	161,69	137,73	88,58	57,50	16,93	0,99	0,95	0,97	0,97	0,96

Table 5.22: Simulated experimental data: Random sampling from Moser-Model

Table 5.23 exhibits the initial conditions and results of the Bayesian model fit and discrimnination of Michaelis-Menten- and Moser-Model with respect to the data in Table 5.22.

Data source	Model	Prior knowledge*					Method	Parameterization						Discrimination	
								Specific value	Parameter						
			μ_{max}	K_S	Y_{XS}	n			μ_{max}	K_S	Y_{XS}	n	RLS	P_{MM}/P_{Moser}
Random sampling Moser	Michaelis-Menten	Parameter	μ_{max}	K_S	Y_{XS}		Bayes	Parameter**	0,50 (0,50)	4,77 (4,77)	0,49 (0,48)			
		Mode	0,58	18,5	0,58			Variance	$4,28*10^{-4}$	11,21	$5,32*10^{-4}$		2,11*10^3	0,23 (0,72)
		Standard dev. *[%]*	50	50	50		RLS	Parameter	0,58	18,5	0,58			
	Moser	Parameter	μ_{max}	K_S	Y_{XS}	n	Bayes	Parameter**	0,48 (0,48)	90,88 (89,81)	0,48 (0,48)	5,08 (4,69)		
		Mode	0,49	94,96	0,52	4,78		Variance	$1,30*10^{-3}$	$2,39*10^{3}$	$7,36*10^{-4}$	4,68	1,36*10^3	
		Standard dev. *[%]*	50	50	50	50	RLS	Parameter	0,49	94,97	0,52	4,78		
* arbitrarly chosen for analytic reasons; mode is based on RLS fit														
** no brackets: posterior mean; in brackets: posterior mode estimate														

Table 5.23: Initial conditions and results of model fit and discrimination towards a random sample from the Moser-Model

Annex F comprises all figures belonging to that simulation experiment.

For the random sample from the Michaelis-Menten-Model the RSL-fit provides a RLS ratio Michaelis-Menten/Moser of 481/501 (Tabel 5.21). If this was the only measure / decision basis for model discrimination this would be just a marginal hint that the Michaelis-Menten-Model (with the referring parameterization from the RLS fit) specifies the true underlying process. The RLS fit just minimizes the summed squares of distances between experimental data and model solution but does not take into consideration which experimental data might be more or less interfered by error in experimentation. The latter is explicitly considered / achieved by the Bayesian method applied here, yielding a relative model probability Michaelis-Menten/Moser of 23,4 (Tabel 5.21) which is a clear proof that in contrast to the RLS-Method the Bayesian-Method is capable of filtering the individual noise from the experimental data to some extent, which yields a clear recovery of the Michaelis-Menten-Model as true underlying process. As prior knowledge the parameterization from the RLS-fit has been applied with a comparatively large standard deviation of 50% of the total parameter value. This represents very little prior knowledge being centered on the region of the real underlying parameterization (which is similar to RLS fit parameterization). By doing so, the influence of the prior knowledge is minimized and maximum priority is placed upon the information contained in the experimental data on model parameterization and probability. It has to be emphasized, that for the parameters μ_{max} and Y_{XS} the Bayesian estimates are close to that of the true underlying values with very little variance (uncertainty) while for parameter K_S the value is much different with comparatively very large uncertainty. This is subject to the low sensitivity of K_S, which means that for this parameter a comparatively large variation shows a comparatively low change in model solutions. Thus a comparatively broad range of parameter values for K_S yields similar solutions, making it hard for the Bayesian method applied here to decide on a narrow spectrum of parameter values with high probability, especially in the presence of noise interfering the experimental data. Figure E.57 and E.58 show, that for both models the simulations with Bayesian and RLS parameterization are very similar. This underlines the interpretation of the parameterization results exposed above (models with comparatively different values for insensitive parameters might have similar solutions). Figure E.59 to E.62 show exemplarily the prior and posterior parameter probability distribution. It has to be emphasized here, how drastically the application of the Bayesian method reduces the uncertainty of parameterization by extracting the information contained in the experimental data. The posterior parameter distribution sample is then further mathematically processed yielding a measure for model validation (Figure E.63 to E.70) and the desired predictions for experimental observations at defined process settings in terms of most probable observation and its uncertainty (variance). Both models

are valid to estimate the true underlying process with respect to the validation criteria. Figure E.71 to E.74 exemplaryly show the probability distribution of experimental observations for t=5 that could theoretically occur if the Michaelis-Menten- and Moser-Model were true. In 90% of all cases (or experiment repetitions at equal conditions) the experimental observations will be located between the 5%- and 95%-quantile. As probability distributions (especially those of Gaussian nature) are theoretically endlessly on both sides (which means that even very extreme values have a certain probability, although extremely small, to be experimentally observed with respect to the error of experimentation / measurement) the range of practical relevant values is reduced to the 90% most probable events. For the Michaelis-Menten-Model the simulated experimental data (which is the interfered solution of the Michaelis-Menten-Model at t=5) is for the substrate concentration right in the center of the distribution. In terms of Bayesian theory and interpretation, it means that this experimental value has (in comparison to the range considered) a comparatively high probability to be experimentally observed. As the occurrence of noise interfering the real underlying value is assumed to be Gaussian distributed with zero noise as expected value, it also means that the observed experimental result is comparatively very close to the true underlying value (if the model was true). The experimental observation for the cell concentration is not that stringently centered in the middle of the distribution compared to the substrate concentration, but is nevertheless still in good support of that distribution (among the 20% of most probable events). Thus this experimental observation can be considered as being still an acceptable estimate of the true underlying value (if the model was true). In general it has to be emphasized that the posterior parameter distribution ascribes to each parameterization in the range considered a certain probability that this parameter is the true parameterization of the real underlying process if the referring model was true. Thus, the range of possible observations between the 5%- and 95%-quantile can be considered as the range that contains by 90% probability the true non-interfered value (if the model was true). If the experimentally observed value is in strong support of that distribution and if the model is valid, than it can be assumed that this value is not significantly interfered by noise and is thus a good estimate of the true non-interfered value. If an experimentally observed value is not in support of the distribution, than it can be assumed, that the value is either extremely interfered by noise (among the 10% of extreme observations at the distribution borders) or the model is a just a poor approximation of the true underlying process (check model validation in that case). According to the Bayesian validation principle exposed in subchapter 2.3.2.2, the Michaelis-Menten-Model is valid with respect to all four validation criteria (Figure E.63 to E.66). For the substrate concentration the validation criteria mean and variance are centered pretty much in the middle of the distribution while for the cell

concentration the validation is less stringent. This observation is in agreement with the location of the experimental cell concentration (at t=5 h) within the posterior distribution of experimental observation (Figure E.71) and might hint to a more precise approximation of the true underlying process by the Michaelis-Menten-Model for the substrate concentration. Due to the Bayesian validation method applied here, also the Moser-Model is a valid approximation of the true underlying process (Figure E.67 to E.70). The location of the experimental validation criteria within the distribution is comparable to that of the Michaelis-Menten-Model and can be interpreted analogues. Also the simulated experimental results at t=5 h (Figure E.73 to E.74) are in good support of the distribution of possible experimental observations, with the cell concentration being more centered. The same interpretation approach as exposed above for the Michaelis-Menten-Model can be applied here. As conclusion it can be stated that both models are a valid approximation of the true underlying process with respect to the Bayesian analysis principles applied here. The simulated experimental data is in all cases in good support of the distribution of values that could be experimentally observed if the referring model was true. The Michaelis-Menten-Model could be clearly recovered as true underlying process by the Bayesian relative model probability. Thus it can be stated with respect to this simulation experiment, that the Bayesian methodology applied here has been implemented correctly and is able to describe and recover the true underlying process (reality) which is known in this artificial experimental setup.

For the random sample from the Moser-Model, the RSL -fit provides a RLS ratio Michaelis-Menten/Moser of 2113/1360. If this is the only measure (decision basis) for model discrimination, this would be a clear proof that the Moser-Model with the referring parameterization from the RLS-fit specifies the true underlying process. The Bayesian method applied here yields a relative model probability of Michaelis-Menten/Moser of 0,7207 which is again a clear proof that the Bayesian-Method is capable of clearly recovering the true underlying process. Again, as prior knowledge the parameterization from the RLS-fit has been applied with a comparatively large standard deviation of 50% of the total parameter value for the reasons already exposed above for the simulated experimental data from the Michaelis-Menten-Model. It has to be emphasized, that for the Moser-Model the posterior variance (uncertainty of prediction) of the parameters μ_{max} and Y_{XS} is very small in comparison to the parameters K_S and n (compare Table 5.23). Here also the interpretation / explanation exposed above for the simulated experimental data from the Michaelis-Menten-Model is valid. The function and interpretation of the distribution of experimental observations is not explained here again, as this has been already done in detail above for the simulated experimental data from the Michaelis-Menten-Model. According to the Bayesian validation principle exposed in subchapter 2.3.2.2, the Moser-

Model is only valid with respect to the cell concentration mean and variance (Figure F.89 to F.92). The same situation can be observed for the validation of the MM-Model (Figure F.85 to F.88), whereas here also the variance of the cell concentration is almost outside the validation range (almost identical with 95%-quantile). These validation results therefore support the conclusion drawn from the relative model probability. Nevertheless, none of these models is valid with respect to the true underlying process (from which this experimental data has been derived) with sufficient probability. This result is caused by the comparatively high noise that has been randomly added to the Moser-Model solutions and, thus, interfers this data. Compared to the random sample from the Michaelis-Menten-Model (Tabel 5.21) also the RLS, which can be considered as measure for the extent of total random noise addition in this case, are significantly higher, which causes the observed differences in validation. As conclusion it can be stated, that the Bayesian method applied here is also in this case able to recover the true underlying model as the most probable one. This is possible here even from data which is significantly more interfered. But also due to that comparatively high interference, both models are rejected as not being valid. These observations again proof the functioning of the method applied, as the model from which the data has been drawn is identified as most probable. But model proposals are also clearly rejected by that Bayesian method, when the real underlying process is superimposed by too much noise. The latter also shows the limit of the method to provide validated models by extracting the information contained in interfered experimental data.

5.2.1.5 How are the results for the real experimental data influenced by the prior variance?

Which results does the method and algorithm implemented here provide for the real experimental data and how does the choice of the prior variance influence these results?

Table 5.24 shows the real experimental data. The true underlying process from which this data originates is unknown. The data is interfered by noise / error in experimentation. The standard deviation is known and has been also applied for random addidition of noise to the simulation experiments (compare Table 5.15).

Data source	Process variable	Data											
	Time [h]	0	1	2	3	4	4,5	5	5,5	6	6,5	7	8
Experiment	Cell conc. [g*L^{-1}]	10	15	24	37	53	70	84	96	101	107	104	111
	Substrate conc. [g*L^{-1}]	200	183	179	146	107	78	51	21	1	0	0	0

Table 5.24: Real experimental data

Table 5.25 exhibits the initial conditions and results for three Bayesian model fit and discrimination approaches towards the real experimental data. Run 1 and 2 are characterized by identical initial conditions (for reasons of reproducibility analysis). Run 3 only differs from Run 1 and 2 in the choice of the prior variance (the prior mean is also identical).

Annex G comprise all figures referring to the model fit, discrimination and analysis approaches of the real experimental data 'Exp. data (Run 1)' and 'Exp. data (Run 3)'. The results for 'Exp. data (Run 2)' are not exposed in the Annex G as they are completely identical to that of 'Exp. data (Run 1)'. The simulation approaches 'Exp. data (Run 1)' and 'Exp. data (Run 2)' additionally provide results for the posterior mean and mode estimate as Bayesian point estimate for model parameterization. The relative model probability is the only result in this analysis that depends on the choice of the point estimate (candidate vector respectively). Therefore, no separate results are exposed in Annex G as they are valid for both approaches.

Data source	Model	Prior knowledge[*]				Method	Parameterization						Discrimination	
							Specific value	Parameter						P_{MM}/P_{Moser}
								μ_{max}	K_S	Y_{XS}	n	RLS	
	Michealis-	Parameter	μ_{max}	K_S	Y_{XS}		Bayes	Parameter[**]	0,46 (0,47)	13,60 (15,08)	0,48 (0,48)			
	Menten	Mode	0,6	18	0,6			Variance	$1,61*10^{-4}$	4,67	$4,02*10^{-4}$		161,91	
Exp. data		Standard dev. *[%]*	50	50	50		RLS	Parameter	0,47	15,99	0,48			5,39 (4,14)
(Run 1)		Parameter	μ_{max}	K_S	Y_{XS}	n	Bayes	Parameter[**]	0,44 (0,43)	23,19 (24,88)	0,47 (0,47)	1,30 (1,31)		
	Moser	Mode	0,6	24	0,6	1,5		Variance	$7,14*10^{-4}$	35,7	$4,29*10^{-4}$	0,02	185,57	
		Standard dev. *[%]*	50	50	50	50	RLS	Parameter	0,42	0,66	0,47	0,81		
	Michealis-	Parameter	μ_{max}	K_S	Y_{XS}		Bayes	Parameter[***]	0,47	15,08	0,48			
	Menten	Mode	0,6	18	0,6			Variance	$1,61*10^{-4}$	4,67	$4,02*10^{-4}$		161,91	
Exp. data		Standard dev. *[%]*	50	50	50		RLS	Parameter	0,47	15,99	0,48			4,29
(Run 2)		Parameter	μ_{max}	K_S	Y_{XS}	n	Bayes	Parameter[***]	0,43	24,88	0,47	1,31		
	Moser	Mode	0,6	24	0,6	1,5		Variance	$7,14*10^{-4}$	35,7	$4,29*10^{-4}$	0,02	185,57	
		Standard dev. *[%]*	50	50	50	50	RLS	Parameter	0,42	0,66	0,47	0,81		
	Michealis-	Parameter	μ_{max}	K_S	Y_{XS}		Bayes	Parameter[**]	0,48 (0,47)	16,26 (15,10)	0,50 (0,48)			
	Menten	Mode	0,6	18	0,6			Variance	$8,95*10^{-5}$	1,77	$3,34*10^{-4}$		161,91	
Exp. data		Standard dev. *[%]*	10	10	10		RLS	Parameter	0,47	15,99	0,48			22,25 (21,34)
(Run 3)		Parameter	μ_{max}	K_S	Y_{XS}	n	Bayes	Parameter[**]	0,44 (0,44)	24,35 (22,63)	0,50 (0,47)	1,31 (1,27)		
	Moser	Mode	0,6	24	0,6	1,5		Variance	$5,48*10^{-4}$	5,52	$5,96*10^{-4}$	0,01	185,57	
		Standard dev. *[%]*	10	10	10	10	RLS	Parameter	0,42	0,66	0,47	0,81		

*mode equals MC start values; variance chosen according to aim of experiment
**no brackets: posterior mean; in brackets: posterior mode estimate
***posterior mode estimate

Table 5.25: Initial conditions and results of model fit and discrimination towards the real experimental data

In the first analysis ('Exp. data (Run 1)' and its repetition 'Exp. data (Run 2)') a standard deviation of 50% of the prior mode has been applied, in the second analysis ('Exp. data (Run 3)') of 10% respectively. The first approach represents significant smaller prior knowledge compared to the second approach, which gives the information contained in the experimental data much more influence on the formation of the posterior distribution. In all approaches, identical prior modes have been chosen that differ significantly from the range assumed for the posterior mode (dimension estimated by RLS-fit). This is done in order to study, to what extent the choice of the prior mode and variance superimposes the information on the model parameterization contained in the experimental data and thus interfers analysis results. For the Moser-Model the prior mode differs more from the assumed posterior range (compare corresponding RLS-parameters, especially for parameters K_S and n). For the analysis with 50% prior standard deviation, the relative model probability (calculated with the posterior mode estimate as candidate vector) is 4,29 (4,14 respectively) which represents a decision for the Michaelis-Menten-Model that is much more clearly than one would assume when only the RLS ratio is considered. This again proofs the advantage of the Bayesian approach towards the RLS fit as already exposed and analyzed in the preceeding chapters. The latter just minimizes the summed squares of distances between experimental data and model solution but does not take into consideration, which experimental data might be more or less interfered by error in experimentation. This is explicitly achieved by the Bayesian method applied here yielding a more distinct decision for the Michaelis-Menten-Model, which is a clear proof that in contrast to the RLS-Method the Bayesian-Method is capable of filtering the individual noise from the experimental data to some extent. The analysis with 10% standard deviation shows a relative model probability of 21,34, which means a significantly more distinct decision for the Michaelis-Menten-Model compared to the result of the analysis with 50% standard deviation. This effect is referred to the fact, that for the Moser-Model the prior mean differs much more from the assumed range of the posterior mode (compared to the Michaelis-Menten-Model). This means, that in case of a declining prior variance (e.g. from 50% to 10% as in this case) the parameterization of the Moser-Model is much more influenced by imprecise prior knowledge and less influenced by the experimental data (compared to the Michaelis-Menten-Model). This automatically leads to an increased relative model probability favoring the Michaelis-Menten-Model. For the analysis with 10% prior variance all validation criteria are fulfilled for the Michaelis-Menten-Model, with the mean and variance of the substrate concentration being in much better support of the distribution compared to the corresponding values of the cell concentration (Figure G.113 to G.116). This can be interpreted as a hint that the model performance is better for the substrate concentration. Although the Michaelis-Menten-Model has been identified as

being much more probable with respect to prior knowledge and exp. data the same observations / conclusions can made for the Moser-Model regarding validation (Figure G.117 to G.120). Thus, the Moser-Model is also a suitable validated approach to simulate the true underlying process. For the analysis with 50% prior variance (Figure G.99 to G.106) the same observations / conclusions can be made as already exposed for the analysis with 10% variance. Thus it can be stated for the experimental data, that both approaches yield validated models with acceptable forecast properties for both the Moser- and the Michaelis-Menten-Model. The choice of prior variance does not affect the model choice but rather affects model averaging due to the magnitude of the relative model probability.

5.2.1.6 Are the results reproducible with acceptable precision?

Are the results of the Bayesian method applied here reproducible with acceptable precision when the method is applied to identical data sets at identical conditions (start values, prior knowledge etc.)? In other words: Is the method correctly tuned (sample sizes, number of Markov-Chains etc.) in order to provide a sufficient reproducibility?

To answer this question, two repetitions of the Bayesian methodology implemented here were applied to the same experimental data set (Table 5.24) at identical conditions. These conditions, as well as the corresponding results, are exposed in Table 5.25 (simulations 'Exp. data (Run 1)' and 'Exp. data (Run 2)'). The corresponding figures are comprised only for the simulation 'Exp. data (Run 1)' in Annex G as the results / figures showed to be identical for both repetitions.

The Bayesian method applied here is based on a Markov-Chain-Monte-Carlo process and other random sampling approaches (Monte-Carlo methods), in which representative samples are drawn randomly from distributions of interest. Due to the nature of random sampling, samples drawn from a distribution can only provide an estimate of a distribution and its characteristic parameters. By nature, a distribution can only be determined with 100% precision when the sample size is infinite, which is computationally not practicable. But already at a certain sample size, the sample can be considered as being representative, as at this size the results calculated from different samples (at same sampling conditions) only differ within an acceptable range. To provide a functional implementation of the Bayesian method that yields reliable results, it is inevitable to tune all sampling processes in order to ensure reproducibility of results with acceptable precision. In this work, I decided to play safe, which means, that all sampling processes are adjusted rather too precise. This requires a lot of computational effort but ensures the generation of reliable results, which in turn also enables to draw reliable conclusions from the data regarding method results and dynamics. The Bayesian method implemented here

has been applied two times to the experimental data at identical conditions (prior knowledge, sample sizes etc.). The mean and variance, which are calculated over the parallel Markov-Chains (at identical step) in defined intervals, are applied as characteristic distribution parameters and indicators of chain convergence. The posterior mode estimate has been taken as candidate vector for calculation of the relative model probability. The results can be taken from Table 5.25 (simulations 'Exp. data (Run 1)' and 'Exp. data (Run 2)'). The posterior mode estimates, as well as the posterior variances, are identical for both runs up to the forth decimal place (for presentation reasons results are only exposed to second decimal place in Table 5.25). The relative model probability of run 1 is only 3,5% lower than that of run 2 (this deviation is not relevant for discrimination / decision). Thus, one can conclude from these results, that the Bayesian method applied here has been properly implemented / tuned in order to provide an acceptable reproducibility of results.

5.2.2 Application of the Bayesian methodology to the real life problem of SPD optimization

In 5.2.1 different simulation studies were achieved for the example of two simple growth models in order to provide evidence for the correct method implementation, as well as to ensure proper understanding of the method dynamics and results interpretation. By this procedure, the initial conditions for an efficient application of the Bayesian methodology to the real life problem of SPD optimization are provided. All conclusions regarding method implementation, adjustment and result interpretation, which have been derived from these initial studies, were applied to the analysis and optimization of SPD. Below, the analysis and optimization of the process variables condenser temperature, evaporator temperature, stirrer rotation and pump power towards the target values rancimat value, acid value and tocopherol concentration is exposed. The analysis is achieved independently for each target value by separate process models. As prior knowledge on model parameterization a Gaussian distribution with RSM-parameterization as mode and a standard deviation of 50% (regarding mode) is applied. For relative model probability estimation both models are considered to exhibit equal prior probability. Finally, the possibilities of simultaneously process optimization towards multiple goals (all target values) are discussed. This is of special importance in case of conflicts regarding modeling and optimization goals.

5.2.2.1 Model fit, discrimination, validation and analysis of rancimat models

In 5.1.2.3 a model of 1st grade, as well as different models of 2nd grade, were derived via RSM. In order to determine significant model terms for the 2nd grade model, three different methods for error variance estimation were applied. For each of these alternatives the significance of model terms was tested by F-test of the relative variance or t-test of the

individual confidence intervals. (These methods are considered as basic statistic knowledge in this work and, thus, are not further exposed in detail.) This results in six alternative methods to derive a candidate for the 2nd grade model. Some of these methods confirm the 1st grade model, while others generate models with significantly more model terms and a higher level of complexity. This results in models of varying complexity. Neither the 1st grade model nor one of the 2nd grade models is valid according to statistic validitiy testing by RSM.

Model	Distribution parameter	Model term															Validation
		β_0	β_1	β_2	β_3	β_4	β_{11}	β_{22}	β_{33}	β_{44}	β_{12}	β_{13}	β_{14}	β_{23}	β_{24}	β_{34}	
1st grade	Mode	9,04		-2,78		1,73									0,99		Yes
	Variance	0,01		0,01		0,01									0,02		
2nd grade	Mode	9,04	0,29	-2,72		1,58	0,41	-0,28	0,56	-0,19					0,86		No
	Variance	0,01	0,02	0,01		0,01	0,02	0,03	0,03	0,01					0,01		

Table 5.26: Bayesian parameterization and validity testing of 1st and 2nd grade model for rancimat

The model proposals exposed in Table 5.26 were applied to Bayesian model fit, discrimination and analysis. Both models were fit to the experimental data by using the Bayesian method implemented in this work. Table 5.26 exhibits for each significant model term (identified by RSM) the mode and variance estimate of the corresponding Bayesian posterior parameter distribution. For the reasons exposed in chapter 5.2.1, the posterior mode estimate is applied to all further modelling and analysis as Bayesian point estimate of model parameterization.

Figure H.121 to H.124 exemplarily show the convergence of the Markov-Chain for the 4th parameter of the 1st grade model. The distribution mean and variance estimates were calculated from the parallel Markov-Chains as characteristic distribution parameters that can be applied as convergence indicators. It can be stated here, that a chain convergence of the Markov-Chain can be observed from the second estimation cycle of the covariance matrix.

Figure H.122 and H.124 show, that from the beginning of the second estimation cycle on, the mean and the variance estimate of the distribution just slightly oscillate in a very small region around the posterior values. These deviations can be explained by the inaccuracies of distribution parameter estimation. The estimation of the posterior distribution was only achievable by the use of random sampling methods which are limited to a sample size that can be properly handled by a computer algorithm. In this situation, the estimates of distribution variance and mean only deviate in the third decimal place. These deviations are negligible with respect to model parameterization and results. The chain convergence for all other parameters of the 1st and 2nd grade model is similar and thus will not be further discussed here or exposed in Annex H.

It must be emphasized, that the 2nd grade model is significantly more probable than the 1st grade model with respect to Bayesian means (Table 5.27). Nevertheless, only the 1st grade model is valid according to the Bayesian validation criteria (Table 5.26, Figure H.130 to H.133).

Probabilities of candidate vector	Log likelihood	Log prior probability	Log posterior ordinate
Model grade 1	-23,3 (-23,3)	-1,62 (-3,6)	-1,59 (-3,6)
Model grade 2	-18,3 (-18,3)	-4,3 (-4,3)	-5,7 (-5,7)
Relative probability $P_{mod\,1}/P_{mod\,2}$	$3,3*10^{-7}$ $(3,3*10^{-7})$		

Table 5.27: Relative model probability estimation of rancimat models

At first sight, this may seem to be a contradiction. However, a more detailed analysis of this issue reveals, that this is a realistic result that might be obtained from a Bayesian model discrimination and analysis.

The applied Gibbs-sampling process yields a posterior distribution sample of parameter vectors that, theoretically, could occur with respect to the experimental data if the model was true (valid). Nevertheless, the latter is an assumption that could potentially be wrong. The mode estimate of the posterior distribution is applied to the relative model probability estimation in the form of a candidate vector (irrespective of model validity). It does explicitly not consider or contain any information regarding the uncertainty of this model parameterization, which is given by the posterior distribution variance. However, if the model is valid, this estimation can be considered as being most probable with respect to likelihood and prior. Equations 2.23 and 2.24 show, that the relevant values for each candidate vector in the relative model probability estimation are the likelihood and the prior-posterior-ratio.

As previously mentioned, the likelihood might exhibit a decisive function in model discrimination by relative model probability estimation. This situation especially occurs in the case of large differences in the number of model terms, where more complex linear models fit the experimental data much better by nature due to a higher flexibility. This results in a larger likelihood of the candidate vector for the more complex 2nd grade model which might be decisive for the discrimination result. It is crucial to mention at this point, that the relative model probability does not determine model validity, but only determines which model would be most probable if both models were true (valid). It should be emphasized, that under this assumption, there is always a most probable model and parameterization by nature irrespective of the grade of deviation from the experimental

data. More complex models are favoured by nature in this context due to a larger likelihood as already exposed above. The aim of modelling is to model the true underlying process and not the experimental data. An experimental data set is merely a snapshot of the true underlying process which is more or less interfered by errors of experimentation. Thus, the question of whether the model is a valid approach to the true underlying process has to be treated separately, irrespective of the fit or likelihood of an experimental data set belonging to that process.

The model validity is determined by the method proposed by Geweke (subchapter 2.3.2.2.5). In this context, functions of the experimental data are applied. The functions chosen and applied here are the mean and variance of the total experimental data set. These functions represent the inner structure of a data set which is less affected or interfered by experimentation errors compared to single data points. With respect to the posterior parameter sample and error of experimentation a posterior probability distribution of function values, which might occur if the model is true and a valid approximation of the true underlying process, can be determined (compare chapter 2.3.2.2.5). This distribution can be viewed as fingerprint of the inner structure of hypothetical experimental data sets that refer to a certain true underlying process.

If the mean and variance of the real experimental data set are in support of the distribution, then the model is valid. Figure H.132 and H.133 show, that the validation method applied here rejects the 2nd grade model as the inner structure of the real experimental data is not in support of the values to be expected if the model was a valid approximation of the true underlying process.

Figure H.130 and H.131 show, that the inner structure of the real experimental data is clearly in support of the posterior distribution of the 1st grade model. In other words, it strongly agrees with the values to be expected if the 1st grade model was true. The effects comprised in the 1st grade model can thus be confirmed as being valid for the true underlying process. It can be also stated that the additional model terms and effects of the 2nd grade model are not confirmed and thus represent classical experimental data overfitting by additional terms. It is overfitting because, on the one hand, these effects or terms improve the fit of a single experimental data set due to higher model flexibility. On the other hand, they adulterate and interfer the recovery and approach of the true underlying process. Thus, the validation method applied here proves to be capable of identifying insignificant model terms (effects) and model overfitting. It must be emphasized here, that the significance of the additional model terms of the 2nd grade model has been questionable anyway since the beginning of the analysis. This model has been confirmed by only one of the six statistical methods applied in the context of model generation via

RSM (compare 5.1.2.3). The Bayesian method applied here proved to be suitable for counterchecking the statistical generation of models by RSM and the corresponding modelling results. Thus, it is an appropriate method that aids in avoiding imprecision of simulations by overfitting. The 1st grade model is also valid with respect to the data set of the orthogonal central composite design as no additional knowledge is provided by this design, compared to the smaller full factorial design, regarding effects of the true underlying process.

This practical example perfectly illustrates the major drawback of conventional curve-fitting and discrimination by RLS in comparison to the Bayesian approach applied here. Minimization of RLS only provides an appropriate fit and modeling of the experimental data and a decision on the model that fits the experimental data the best. In contrast, the Bayesian approach identifies the model and parameterization that is most probable and valid with respect to the true underlying process to be simulated.

For the rancimat value, a 1st grade model derived by RSM (determination of significant model terms by method 1/a) and a 2nd grade model also derived by RSM (determination of significant parameters by method 2/b) have been parameterized by Bayesian Markov-Chain-Monte-Carlo, the Metropolis-Hastings-Algorithm respectively, and discriminated by the application of the Chib and Jeliazkov's proposed method (Chib, S. and Jeliazkov, I., 2001, pp. 270-281). A relative model probability of 1st grade to 2nd grade model of $3,3*10^{-7}$ was computed.

This extremely clear decision for the model of 2nd grade leads to the question of whether this value is realistic in the context of the Bayesian method applied here or just caused by deficiencies of this method regarding the discrimination of models that differ in the numbers of parameters. If the method is adequate regarding this dimensionality aspect, then the question arises by which mechanisms / dynamics such extreme relative probability values are obtained.

During the Bayesian parameterization process for each model by Gibbs-sampling, no problems can arise due to differences in the dimension of the prior parameter distribution. This is subject to the fact that during the parameterization process each model is exclusively analyzed and compared without interactions with the competing model. The posterior parameter sample is exclusively formed by the comparison of parameter vectors of the same model with respect to likelihood and prior probability (compare chapter 2.3.2.2.2 and Eq. 2.21). Thus, the prior probabilities, which influence the probability of accepting a proposed new state given a current state, are derived from the same distribution and therefore have the same dimension. Furthermore, also model validation and the calculation of the posterior distribution of experimental observations at defined

process settings are not subject to dimensionality differences in model parameterizations. This is due to the fact that each model is exclusively analyzed here without any interactions between the competing alternative models.

When models are discriminated by the application of the method proposed by Chib and Jeliazkov, the relative model probability is calculated as the ratio of the marginal likelihood estimates in accordance with Eq. 2.23 and 2.24 (Chib, S. and Jeliazkov, I., 2001, pp. 270-281). In principal, three different characteristic values for each model are set into relation here: The posterior ordinate, the likelihood and the prior probability of the candidate vector. Equation 2.23 and 2.24 show, that the relative model probability can be interpreted as the product of the likelihood ratio and the ratio of the quotients of prior probability and posterior ordinates. With regards to the likelihood, it is assumed that no problems arise from differences in the number of model parameters. This is true because, irrespective of parameterization, each model yields solutions for the same target value (e.g., rancimat value) at identical process variable settings which are compared to the same experimental data in terms of likelihood values. The likelihood is the probability of experimental data occurrence given a defined model and error of experimentation (Eq. 2.19). Thus, the comparison of likelihoods that refer to models with different number of parameters is valid here. Even though the candidate vectors and their prior and posterior probabilities differ significantly in dimension, this does not affect or interfere the relative model probability estimate because no single probabilities for the candidate vectors are set into relation but quotients of prior and posterior probability. Through this standardization and normalization, comparable quantities of equal dimension are applied. These quantities therefore avoid the interference of relative model probability estimates due to differences in dimensionality. Thus, also the comparison of prior and posterior probabilities of candidate vectors with different dimension is valid here.

This assessment has been practically proven by two different model discrimination approaches. In the first approach the model discrimination and corresponding algorithm is applied to the original 1st grade model with four parameters and the 2nd grade model with nine parameters. In the second approach the model of 1st grade is simulated based on the model of 2nd grade. This is achieved by applying the model structure of the 2nd grade model to the 1st grade model and then setting the parameter values of the additional model terms to zero. Thus, formally and technically models of same structure and parameter vectors of same dimensions are compared. Table 5.27 shows, that both approaches (results of second approach in brackets) yield the same results except for that of the prior and posterior ordinates of the candidate vector which are subject to the dimension of the parameter vector. Nevertheless, these differences are balanced out by the application of probability ratios as exposed above. The ratios of the prior and posterior probability

(posterior ordinate) for the candidate vectors are nearly equal, with only negligible differences caused by the random sampling methods applied. The results in Table 5.27 correspond to Eq. 2.23 and 2.24 and demonstrate that the higher probability of the 2nd grade model of up to 7 decimal powers is primarily caused by differences in the likelihoods of the candidate vectors.

Exp. setting		1	2	3	4	5	6	7	8	9	10	11	12	13
Rancimat *[h]*		8,725	3,765	9,78	3,895	12,825	11,815	12,45	11,51	7,485	3,75	9,315	3,855	11,73
Log likelihood of	Model grade 1	-0,45	-0,43	-0,85	-0,49	-0,45	-0,78	-0,41	-0,53	-2,01	-0,43	-0,48	-0,47	-0,9
candidate vector	Model grade 2	-0,53	-0,58	-0,69	-0,5	-0,4	-0,51	-0,53	-0,41	-1,25	-0,4	-0,79	-0,42	-0,63
Exp. setting		14	15	16	17	18	19	20	21	22	23	24	25	
Rancimat *[h]*		10,985	12,535	11,165	10,09	9,715	4,18	11,98	9,745	9,945	11,265	3,76	9,98	
Log likelihood of	Model grade 1	-0,41	-0,4	-0,4	-1,18	-0,72	-1,03	-1,11	-0,75	-0,98	-0,43	-6,2	-1,03	
candidate vector	Model grade 2	-0,41	-0,44	-0,48	-0,43	-0,71	-0,4	-0,4	-0,4	-0,42	-0,83	-4,02	-1,71	

Table 5.28: Likelihood values of exp. data for 1st and 2nd grade model of rancimat value

The likelihood of a data set is the product of the likelihoods which refer to the experimental results of the individual process settings. Table 5.28 shows the extent to which the individual experimental settings contribute to the overall likelihood value. The most significant experimental settings in that context are indicated in red. It is obvious, that the additional experimental settings of the orthogonal CCD (setting 17 to 25) are most important in that context. These settings are essential for the identification of the additional significant model terms of the 2nd grade model, which contribute significantly to the higher likelihood of the respective candidate vector due to the improved overall flexibility and, thus, experimental data fit. Table 5.27, Eq. 2.23 and Eq. 2.24 show, that (besides the likelihood) the ratio between prior and posterior probability of the candidate vectors exhibits a minor contribution to the higher probability of the 2nd grade model (2 decimal powers). For the 2nd grade model, the posterior probability is significantly lower compared to the prior probability which results in a higher value for the prior-posterior-ratio of the 2nd grade model candidate vector. For the 1st grade model, however, the values are nearly equal.

	Optimum setting				Max. rancimat *[h]*	
Process variable	X_1	X_2	X_3	X_4	Simulation	Experiment
Model grade 1	Free	-1	Free	1	12,56	12,54
Model grade 2	1	-1	1	1	12,87	Not valid

Table 5.29: Optimum process settings for rancimat value (Bayesian model parameterization)

Table 5.29 shows the optimum process setting (within the range of analysis) for a maximized rancimat value. Figure H.128 exposes the posterior distribution of hypothetical experimental observations at optimum setting for the 1st grade model. The actual experimental result is in perfect support of the distribution, which underlines the validity and accuracy of that modeling approach. The quality criterion of minimum 8 h is fulfilled

within the whole range of relevant possible events (between 5%- and 95%-quantile) that might occur at optimum process setting if the model was true (valid). Thus it can be stated, that there is only a negligible chance that the quality criterion is not fulfilled at the optimum process setting calculated from the 1st grade model. The graphical exposition of the 1st grade model, given by Figure H.125, helps to identify areas of process settings that might fulfill the rancimat quality criterion and also other optimization goals. As can be seen in this figure, there is also a broad range of other process variable settings which fulfill this criterion. This is of special importance for optimization towards multiple objectives that are characterized by conflicts regarding process adjustments.

5.2.2.2 Model fit, discrimination, validation and analysis of acid value

In 5.1.2.4 models of 1st and 2nd grade with statistically significant model terms were derived by RSM. The models exposed in Table 5.30 were selected to be parameterized, discriminated and analyzed by Bayesian means here. The Figures regarding Markov-Chain convergence, parameter distributions, distributions of experimental observations and validation are comprised in Annex I.

Figure I.134 to I.137 exemplarly show for the 1st grade model the Markov-Chain convergence for one model parameter. The convergence dynamics are similar for all other model parameters. Therefore they are not exposed in that work for reasons of space. A stable convergence of the distribution parameters towards a boundary posterior value can be clearly observed. The chain oscilliates around that boundary value already from update number 3 on. Therefore it is valid to take all chain steps from update number 10 on as representative sample of the posterior distribution. In this case it is obvious, that the computational effort, which is necccessary to gain a representative sample of sufficient size, can be reduced significantly here. It can be also observed here, that in some cases the mean of the Markov-Chains does not approach the mode estimate of the posterior distribution. In this case, the posterior distribution is not symmetric with respect to the distribution expected value, which therefore differs from the mode.

Model	Distribution parameter	Model term															Validation
		β_0	β_1	β_2	β_3	β_4	β_{11}	β_{22}	β_{33}	β_{44}	β_{12}	β_{13}	β_{14}	β_{23}	β_{24}	β_{34}	
1st grade	Mode	0,49	0,11	-0,26		0,12					-0,11				-0,15		Yes
	Variance	$6,56 \cdot 10^{-4}$	$6,93 \cdot 10^{-4}$	$8,01 \cdot 10^{-4}$		$7,33 \cdot 10^{-4}$					$8,02 \cdot 10^{-4}$				$8,83 \cdot 10^{-4}$		
2nd grade	Mode	0,49	0,12	-0,25		0,11		0,16			-0,12				-0,14		Yes
	Variance	$6,71 \cdot 10^{-4}$	$6,64 \cdot 10^{-4}$	$8,05 \cdot 10^{-4}$		$6,85 \cdot 10^{-4}$		$1,6 \cdot 10^{-3}$			$8,16 \cdot 10^{-4}$				$8,92 \cdot 10^{-4}$		

Table 5.30: Bayesian parameterization and validity testing of 1st and 2nd grade model for acid value

Figure I.140 and I.141 expose the prior and posterior parameter distribution exemplary for one parameter of the 1st and 2nd grade model. It has to be emphasized, how drastically the Bayesian method applied here reduces the uncertainty (variance) of predictions by extracting the information on model parameterization contained in the experimental data.

The figures are similar for all other model parameters and are therefore not exposed in that work due to space reasons.

Figure I.144 to I.147 show for both models and validation criteria that the experimental values are in very good support of the posterior distribution. Thus, both models are validated beyond doubt with respect to the true underlying process to be modeled. It has to be emphasized here, that also the 1st grade model has been fit and analyzed regarding its performance on the results of the orthogonal CCD. Therefore it can be stated here, that by means of Bayesian analysis the 1st grade model is sufficient to model the true underlying process. Although the 2nd grade model provides more accuracy due to an additional model term, this effect therefore can be considered as being not crucial for an appropriate modeling of the true underlying process. Thus, the additional experimental effort of the orthogonal CCD, which is necessary for identification of that model term / effect, is not essential in that contex and could be saved in optimization approaches of scaled-up and / or similar SPD processes. The squared effect of the evaporator temperature on the acid value is statistically significant but plays a minor role compared to the other various linear effects and interactions of the process variables on this target value.

Probabilities of candidate vector	Log likelihood	Log prior probability	Log posterior ordinate
Model grade 1	-15,89	-2,43	-2,42
Model grade 2	-13,27	-2,80	-2,82
Relative probability $P_{mod\ 1}/P_{mod\ 2}$	$2,2*10^{-3}$		

Table 5.31: Relative model probability estimation of acid value models

Exp. setting		1	2	3	4	5	6	7	8	9	10	11	12	13
Acid value [mg KOH*g⁻¹]		0,30	0,31	0,27	0,27	1,19	0,56	1,57	0,66	0,29	0,22	0,22	0,32	0,62
Log likelihood of	Model grade 1	-0,52	-0,43	-0,45	-0,40	-0,42	-0,60	-1,84	-0,41	-0,50	-0,42	-0,41	-0,45	-0,82
candidate vector	Model grade 2	-0,45	-0,40	-0,41	-0,41	-0,46	-0,89	-1,60	-0,52	-0,43	-0,48	-0,40	-0,40	-0,81
Exp. setting		14	15	16	17	18	19	20	21	22	23	24	25	
Acid value [mg KOH*g⁻¹]		0,36	0,83	0,30	0,43	0,25	0,29	1,10	0,40	0,34	0,43	0,38	0,33	
Log likelihood of	Model grade 1	-0,54	-0,41	-0,42	-1,00	-0,49	-0,73	-1,23	-0,49	-0,67	-1,09	-0,45	-0,71	
candidate vector	Model grade 2	-0,45	-0,41	-0,40	-0,54	-0,43	-0,42	-0,45	-0,42	-0,41	-0,51	-0,76	-0,41	

Table 5.32: Likelihood values of exp. data for model of 1st and 2nd grade acid value

Table 5.31 shows, that according to the Bayesian relative model probability the 2nd grade model is significantly more probable. This is no contradiction to the fact, that both models are valid, as the latter evaluates the general suitability of a model irrespective of competing models. A more complex 2nd grade model exhibits a higher probability by nature as it approaches the experimental data much better due to a higher flexibility. This aspect is represented by the likelihood value, which is much larger for the 2nd grade model and

contributes almost exclusively to the difference in model probability. The ratio of prior probability and posterior ordinate is almost equal for both models (compare Tabel 5.31). By that observation, again, the special importance of the candidate vector choice and likelihood in relative model probability estimation is underlined. Table 5.32 exposes the contribution of the single process settings to the overall likelihood of the experimental data set with respect to the candidate vector. Again, it is the additional process settings of the orthogonal CCD that decisively contribute to the likelihood differences of the candidate vectors. These additional process settings are crucial for the identification of the additional squared effects of the 2nd grade model. It provides a better fit of the total orthogonal CCD data and thus a higher likelihood of this data set with respect to the candidate vector.

	Optimum setting				Min. acid value *[mg KOH*g^{-1}]*	
Process variable	X_1	X_2	X_3	X_4	Simulation	Experiment
Model grade 1	-1	1	free	1	0,1966	0,25
Model grade 2	-1	-0,0174	free	-1	0,1365	0,2

Table 5.33: Optimum process settings for acid value (Bayesian model parameterization)

Table 5.33 exhibits the process settings (within the range of analysis), which minimize the acid value with respect to the 1st and 2nd grade model. For these process settings also the simulation and experimental results are exposed. It can be stated here, that the quality requirements of at least 0,5 mg KOH*g^{-1} can be realized by the SPD at these process settings. As the acid values, which are achievable at the process settings exposed in Table 5.33, are much lower compared to the quality requirement, it can be assumed, that there is also a broad range of alternative process settings at which the quality requirement for the acid value is fulfilled. This is of special importance for optimization towards multiple objectives which are characterized by conflicts regarding process variable adjustment. The existence of alternative process settings for each objective allows for compromises in optimization towards multiple objectives.

Figure I.142 and I.143 show the posterior distribution of experimental observations at the optimum process setting for the 1st and 2nd grade model. The experimental values (compare Table 5.33) are in very good support of that distribution, which again underlines model validity and prediction accuracy. It also has to be mentioned here, that the entire relevant range of possibly occurring events (between 5%- and 95%-quantile) agrees with the quality criterion to be fulfilled. Therefore, there is only a negligible possibility of missing the quality criterion at the simulated optimum process setting.

Figure I.138 and I.139 graphically expose the dynamics of the 1st and 2nd grade model (for a constant optimum pump power). Such graphics aid the identification of process settings in process optimization towards multiple goals.

5.2.2.3 Model fit, discrimination, validation and analysis of tocopherols

In 5.1.2.5 models of 1st and 2nd grade with statistically significant model terms were derived by RSM. The models exposed in Table 5.34 were selected to be parameterized, discriminated, validated and analyzed by Bayesian means in this chapter. The Figures regarding Markov-Chain convergence, parameter distributions, distributions of experimental observations and validation are comprised in Annex J.

Figure J.148 to J.151 exemplaryly show for the 1st grade model the Markov-Chain convergence for one model parameter. The convergence dynamics are similar for all other model parameters. Therefore they are not exposed in that work for reasons of space. A stable convergence of the distribution parameters towards a boundary posterior value can be clearly observed. The chain oscilliates around that boundary value already from update number 2 on. Therefore it is valid to take all chain states from update number 10 on as representative sample of the posterior distribution (as done in this work). In this case it is obvious, that the computational effort, which is necccessary to gain a representative sample of sufficient size, can be reduced significantly here

Model	Distribution parameter	Model term															Validation
		$ß_0$	$ß_1$	$ß_2$	$ß_3$	$ß_4$	$ß_{11}$	$ß_{22}$	$ß_{33}$	$ß_{44}$	$ß_{12}$	$ß_{13}$	$ß_{14}$	$ß_{23}$	$ß_{24}$	$ß_{34}$	
1st grade	Mode	22,55		-14,92	-2,25	6,47											Yes
	Variance	0,11		0,14	0,13	0,14											
2nd grade	Mode	22,50		-14,99	-2,23	6,53		2,05			-1,56						Yes
	Variance	0,11		0,14	0,13	0,14		0,27			0,13						

Table 5.34: Bayesian parameterization and validity testing of 1st and 2nd grade models for tocopherol

Figure J.154 and J.155 expose the prior and posterior parameter distribution exemplary for one parameter of the 1st and 2nd grade model. It has to be emphasized, how drastically the Bayesian method applied here reduces the uncertainty (variance) of predictions by extracting the information on model parameterization contained in the experimental data. The figures are similar for all other model parameters and are therefore not exposed in that work due to space reasons.

Figure J.158 to J.161 show for both models and validation criteria that the experimental values are in very good support of the posterior distribution. Thus, both models are validated beyond doubt with respect to the true underlying process to be modeled. It has to be emphasized here, that also the 1st grade model has been fit and analyzed regarding its performance on the orthogonal CCD data. Therefore it can be stated here, that by means of Bayesian analysis the 1st grade model is sufficient to model the true underlying process. Although the 2nd grade model provides more accuracy due to two additional model terms, this effect and interection can be considered as being not crucial for an appropriate modeling of

the true underlying process. Thus, the additional experimental effort of the orthogonal CCD, which is necessary for identification of that model terms, is not essential in that contex and could be saved in modeling and optimization of similar and / or scaled-up SPD processes. The squared effect of the evaporator temperature on tocopherol, as well as the interaction between condenser and evaporator temperature, are statistically significant but play a minor role compared to the other various effects of the process variables on this target value.

Probabilities of candidate vector	Log likelihood	Log prior probability	Log posterior ordinate
Model grade 1	-38,48	-1,65	-1,61
Model grade 2	-33,31	-2,40	-2,43
Relative probability $P_{mod\ 1}/P_{mod\ 2}$	$5,80*10^{-6}$		

Table 5.35: Relative model probability estimation of tocopherol models

Exp. setting		1	2	3	4	5	6	7	8	9	10	11	12	13
Tocopherol $[mg*100g^{-1}]$		10,92	1,12	15,67	1,08	44,01	32,22	46,78	36,12	8,94	0,96	16,06	15,19	44,19
Log likelihood of	Model grade 1	-0,47	-0,77	-0,43	-0,82	-0,81	-1,33	-0,43	-1,03	-1,05	-0,72	-0,41	-11,1	-0,88
candidate vector	Model grade 2	-0,4	-1,35	-0,42	-0,47	-0,4	-0,58	-0,56	-0,47	-2,2	-0,4	-0,76	-8,11	-1,37
Exp. setting		14	15	16	17	18	19	20	21	22	23	24	25	
Tocopherol $[mg*100g^{-1}]$		29,11	44,05	31,02	23,86	18,63	0,99	44,08	17,11	24,51	34,1	1,82	21,29	
Log likelihood of	Model grade 1	-0,41	-0,75	-0,78	-0,53	-1,58	-0,42	-0,41	-0,79	-0,51	-0,84	-10,7	-0,52	
candidate vector	Model grade 2	-0,58	-0,48	-0,48	-1,09	-0,78	-0,99	-0,73	-0,43	-0,42	-1,63	-7,79	-0,41	

Table 5.36: Likelihood values of exp. data for model of 1st and 2nd grade tocopherol

Table 5.35 shows, that by means of Bayesian relative model probability estimation, the 2nd grade model is significantly more probable. This is no contradiction to the fact that both models are valid as the latter evaluates the general suitability of a model irrespective of competing models. A more complex 2nd grade model exhibits a higher probability by nature, as it approaches the experimental data much better due to a higher flexibility. This aspect is represented by the likelihood value which is much larger for the 2nd grade model and contributes almost exclusively to the difference in model probability (the ratio of prior probability and posterior ordinate is almost equal for both models). By that observation, again, the special importance of the candidate vector choice and likelihood in relative model probability estimation is underlined. Table 5.36 exposes the contribution of the single process settings to the overall likelihood of the experimental data set with respect to the candidate vector. It can be observed here, that experimental settings from all parts of the orthogonal CCD contribute to the higher likelihood of the 2nd grade model.

Process variable	Optimum setting				Max. tocopherol $[mg*100g^{-1}]$	
	X_1	X_2	X_3	X_4	Simulation	Experiment
Model grade 1	Free	-1	-1	1	46,19	45,42
Model grade 2	1	-1	-1	1	48,22	46,78

Table 5.37: Optimum process settings for tocopherol (Bayesian model parameterization)

Table 5.37 exhibits the process settings (within the range of analysis) which maximize the tocopherol content with respect to the 1st and 2nd grade model. For these process settings also the simulation and experimental results are exposed. The tocopherol content should be as high as possible due to its special nutritional and antioxidant properties. It can be assumed, that there is a broad range of alternative process settings at which acceptable tocopherol contents can be realized. This is of special importance for optimization towards multiple objectives which are characterized by conflicts regarding process variable adjustment. The existence of alternative process settings for each objective allows for compromises in process optimization towards multiple goals.

Figure J.156 and J.157 show the posterior distribution of experimental observations at the optimum process setting for the 1st and 2nd grade model. The experimental values (compare Table 5.37) are in very good support of that distribution which, again, underlines model validity and prediction accuracy.

Figure J.152 and J.153 graphically expose the dynamics of the 1st grade model (for a constant optimum pump power) and the 2nd grade model (for a constant optimum stirrer rotation). Such graphics aid the identification of process settings in process optimization towards multiple goals.

5.2.2.4 Conclusions for the process optimization towards multiple target values

Table 5.29, 5.33 and 5.37 exhibt the process settings at which an optimization (maximization or minimization) of the respective target values can be realized. For the rancimat and acid value, the stirrer rotation is a statistically insignificant process variable, which therefore can be adjusted to the minimum value to increase the tocopherol content. The adjustment of the condenser temperature is (within the range of analysis) not statistically relevant for the rancimat value and exhibits (according to the 2nd grade model) only a comparatively small influence on the tocopherol content by interaction with the evaporator temperature (compare corresponding model term in Table 5.34). Regarding this interaction, the change in the tocopherol content, which is caused by a switch of the condenser temperature from minimum to maximum (and vice versa), is considered as being negligible with respect to the dimension of the totel tocopherol content and the fact, that there are no mandatory threshold levels for this quality parameter. Thus, the condenser temperature can be minimized to serve the important objective of acid value minimization (mandatory quality parameter with threshold level at 0,5 mg $KOH*g^{-1}$). The condenser temperature exhibits the same dynamics in the 1st

and 2nd grade model for the acid value. Its lowering also lowers the acid value. In contrast to that, these models yield an opposite optimum setting for the pump power, which is subject to the fact, that the 2nd grade model includes an additional quadratic effect of the evaporator temperature, which influences also the optimum setting of other process variables due to complex interactions. As far as the adjustment of the pump power is concerned, a conflict in objectives between the acid value minimization (compare 2nd grade model) on the one hand and the rancimat value and tocopherol content on the other hand can be observed. Decreasing the pump power positively affects the acid value but negatively affects tocopherol content and rancimat value. The opposite situation occurs for the evaporator temperature. Thus, the adjustment of pump power and evaporator temperature has to compromise between different objectives. The exact choice has to be subject to further analysis with respect to the individual needs and requirements of the SPD operator or customer, such as individual quality requirements (final application of the product), energetic aspects and process duration.

For the actual situation and process at hand, the pump power would be adjusted to the mean of the analysis range and the evaporator temperature to the acid value optimum. At this adjustment, an acid value of 0,41 mg KOH*g^{-1} is predicted by the 2nd grade model (below the threshold level of 0,5 mg KOH*g^{-1}) as well as a rancimat value of 9,1 h by the 1st grade model (above threshold level of 8 h). An adjustment of the pump power to the maximum value would optimize the rancimat value but would also lead to an acid value of 0,52 mg KOH*g^{-1} according to the 2nd grade model (above the threshold value).

The opposite situation would occur for the adjustment of the pump power to the minimum. On the one hand this situation would optimize the acid value but on the other hand would also result in a rancimat value of 7,4 h according to the 1st grade model (evaporator temperature adjusted to the optimum for the acid value as mentioned above).

6 Summary and Outlook

6.1 Summary

According to the actual state of research, the heat induced formation of 3-MCPD-FE and related substances during the deodorization process of edible oils is a seriouse problem due to significant carcinogenic and mutagenic properties of these substances. This is especially true for refined palm oil, which exhibits critical concentrations and is also of special economic importance for the worldwide food industry.

Actually, the optimization and modification of the refining process is the only alternative to reduce the contaminant formation at present, as the control of the precursor concentrations (e.g. chlorinated residues of pesticides, mono- and diglycerides) during the crude oil production is not possible yet due to divergent economic and political interests of the producers (main producing countries, respectively). Among the different stages of the refining process the modification and optimization of the deodorization process is the most promising approach to lower the heat induced formation of 3-MCPD-FE and related substances below the recommended threshold value.

In this work, the Short Path Distillation (SPD) proved to be a sophisticated and highly effective deodorization technology, which is capable of coping with the heat induced contaminant formation while ensuring the basic quality standards of refined palm oil. Actually, the contaminant formation could be reduced close to the limit of detection, while ensuring the threshold levels for acid and rancimat value. This could be realized along with a comparatively high total tocopherol content in the refined edible palm oil. A major reason for that capability are the special technological characteristics of SPD, which are the operation at fine vacuum, the evaporation of volatile compounds from a thin and constantly renewed liquid film, as well as the efficient separation of these volatile compounds via a short path to the axially arranged condenser unit.

The Response Surface Methodology (RSM) is a sophisticated statistic approach to derive linear process models for process optimization from a set of mathematically designed experiments (SDoE) which maximize the information on the true underlying process at a minimum of experimental effort (costs). This method could be successfully implemented and studied for a literature example in a first step. Afterwards it could be successfully transfered and applied to the real life problem of SPD optimization. Statistically significant models could be identified for the major oil quality criteria acid value, rancimat value and tocopherol content. The measured concentrations of 3-MCPD and related substances were mostly close to the limit of detection or within the standard deviation of experimentation. Thus, a further model based minimization was not possible and necessary here.

The Bayesian methodology for model fit, discrimination, validation and analysis could be successfully implemented and studied for a simple growth model example. This approach

110

ensured, that the special features of this methodology, as well as the interpretation of results and dynamics, are correctly understood and thus could be efficiently applied to the real life problem of SPD optimization. Consequently, these initial studies (example growth model) were logically structured into six basic questions / problems (compare chapter 3.2.1 and 5.2.1), which are essential for a proper method validation and application.

It can be stated here, that due to the successful accomplishment of these initial tasks a proper basis for an efficient application of the methodology to the real life problem of SPD optimization could be provided. Thus, a set of validated process models could be provided, which enable the identification of process settings at which single and multiple objectives of process optimization could be successfully achieved and experimentally proofed.

Additionally, the Bayesian methodology provided important probibablistic measures for the decision among model alternatives and the uncertainty of results / predictions. These features especially enable the evaluation and reduction of risks and costs. Therefore it can be stated in summary, that this thesis provides a coherent stochastic approach of model generation (RSM) and analysis (Bayesian methodology) by which a comprehensive analysis, assessment and optimization of the SPD process could be achieved. By these means, mathematical evidence for the suitability of the SPD to control contaminant formation, while ensuring other major quality parameters, could be provided.

Finally a program / algorithm could be developed, that also enables a convenient application of this coherent stochastic process optimization method to any kind of other technological processes.

6.2 Outlook

In the near future the awareness of the consumer for healthy foods will further increase due to a global trend towards a sustainable production of healthy foods and an increased focus of science, politics and media on this topic.

This development is additionally enforced by an increasing political and legal pressure in this context. Thus it is expected, that the application of innovative and mild refining methods for the production of healthy and contaminent free edible plant oils will be intensified by the producers. This is especially true for palm oil, as it is indispensable for the world wide food industry due to its unique technological properties and crop yield, but also exhibits a high potential of contaminant formation at conventional deodorization conditions.

The SPD proved to be the most suitable method in this context. Actually, an industrial large scale application of this technology is primaryly limited to high-value products, such as pharmaceuticals, due to the high investment costs. Nevertheless, it is expected for the near future that also the production of refined bulk products, such as palm oil, will switch over to SPD due to a significantly increasing public, political and legal pressure regarding health and quality standards. Also increasing energy costs, as well as an increasing awarenesss for

energy consumption and climate change (concomitant with a corresponding legislation), will force the refining industry to switch over from highly energy consuming conventional deodorization processes to SPD. Thus, a coherent model based analysis and optimization of SPD is a future-oriented approach that will gain increasingly more scientific attention.

In general, model-based process and product optimization approaches will gain more attention and importance also in the industrial food production. A major reason is the increasing global competition and an increasing consumer demand for cheap and high-quality products. This inevitably leads to an increased demand for a systematical understanding of inner process dynamics and a precise identification of process settings that optimize costs and quality. These demands cannot be fulfilled by trial-and-error approaches anymore, but require coherent and self-contained mathematical approaches that provide a maximum of information at a minimum of experimental effort (costs).

The method of choice in this context is the application of RSM (SDoE, respectively), as it is a so-called black-box approach that does not require a deeper understanding of physico-chemical processes. At the moment a major drawback for a widely and routinely use of that method in the food industry is the simple fact, that there is still little affinity to mathematical R&D approaches in that industrial sector. But this situation will be inevitably overcome by the simple need to be competitive regarding prices and quality in a globalized market. Thus, it can be stated here, that it is quite a future-oriented approach to treat product and process optimization problems in the food production by means of RSM (SDoE, respectively).

The Bayesian approach towards model fit, discrimination and analysis is highly sophisticated, as it provides an estimation of the true underlying process with respect to the error of experimentation, as well as uncertainty measures of results and predictions. This enables a coherent assessment of risks and costs, which is inevitable for a modern R&D-management.

In contrast, conventional approaches (such as RLS-methods) only provide modelling of the experimental data, which might significantly diverge from the true underlying process due to experimentation noise (especially in case of few repetitions of experimental settings). Thus, Bayesian analysis can be considered as a highly efficient process and product optimization tool even in the R&D of the food industry.

This is especially true with respect to the above mentioned future requirements of food R&D and production in a globalized market. Nevertheless, a major drawback for a standard application is the high mathematical complexity, as well as the high computational effort. This is especially true for the food industry for the reasons mentioned above. To ease the practical application, future research should focus on the design of easily handable tools that do not require deeper mathematical knowledge, as well as on the reduction of the computational effort (time consumption). The latter is basically caused by the application of Markov-Chain-Monte-Carlo approaches, which are based on random sampling / random processes. In this context the method performance depends on the sample size, which in

turn determins the computation time. Therefore future research should also focus on the acceleration of Markov-Chain convergence and the generation of more representative samples at lower sample sizes.

In this thesis process modeling and optimization was achieved based on the model with the higher probability. Nevertheless all competing models exhibit a certain probability of being true. This fact is considered in Bayesian model averaging approaches that generate a concerted simulation, which is proportionally mixed according to the relative probability of the competing models. Further research should therefore also focus on the implementation of a Bayesian model averaging approach, and evaluation of the same, with respect to the approach applied in this work. In this context it is of special interest, if the improved preciseness justifies the additional computational and programming effort.

Annex A – Pictures of SPD process and plant
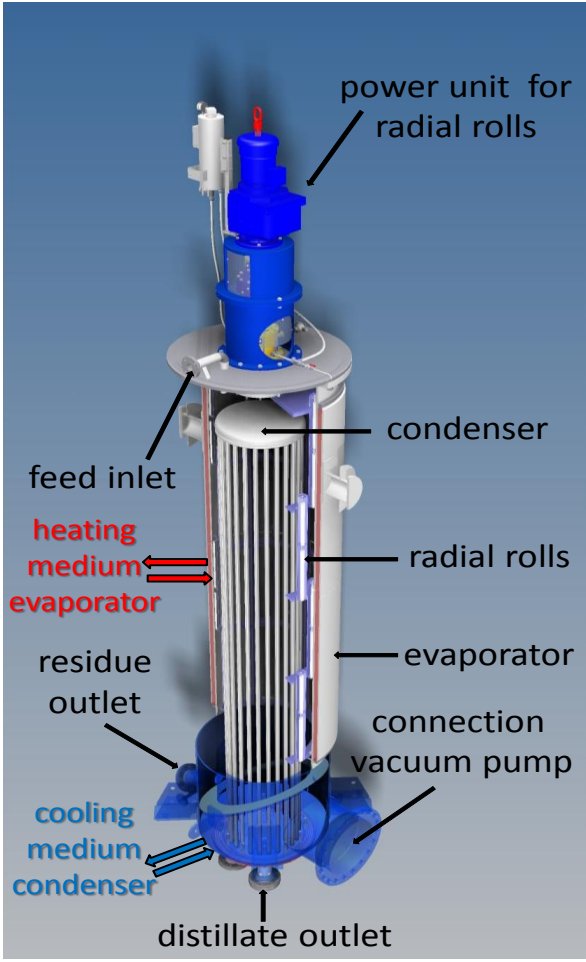
Figure A.1: Basic design of SPD evaporator unit with integrated condenser

Figure A.2: Technical scale SPD evaporator unit (glas) with integrated spiral condenser
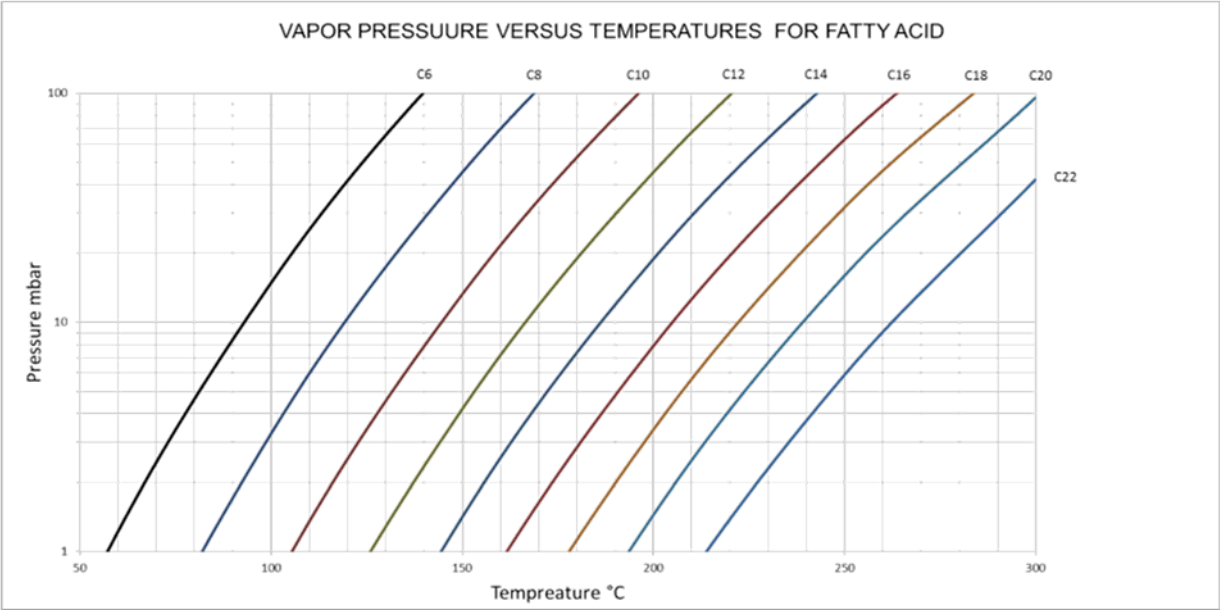
Figure A.3: Vapor pressure curves of fatty acids
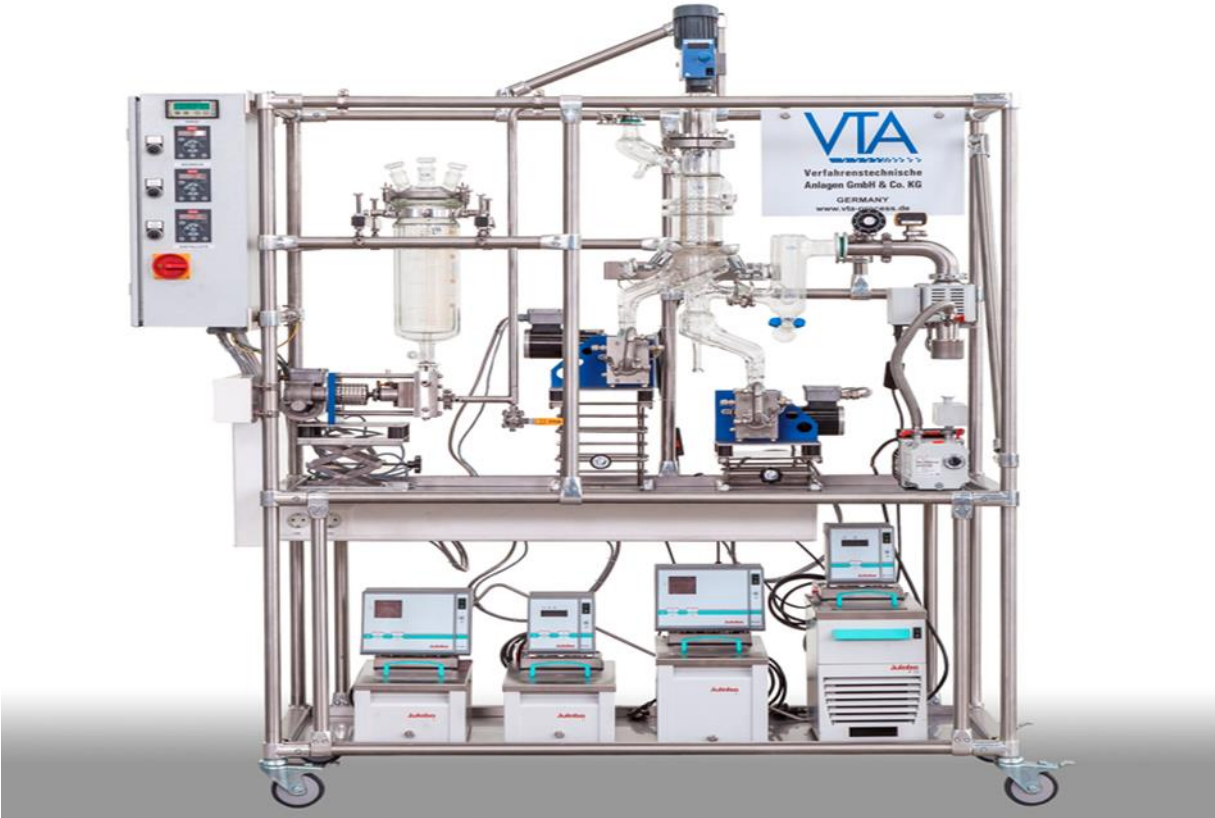
Figure A.4: Setup of SPD plant VKL-70-5 (VTA GmbH, Niederwinkling)
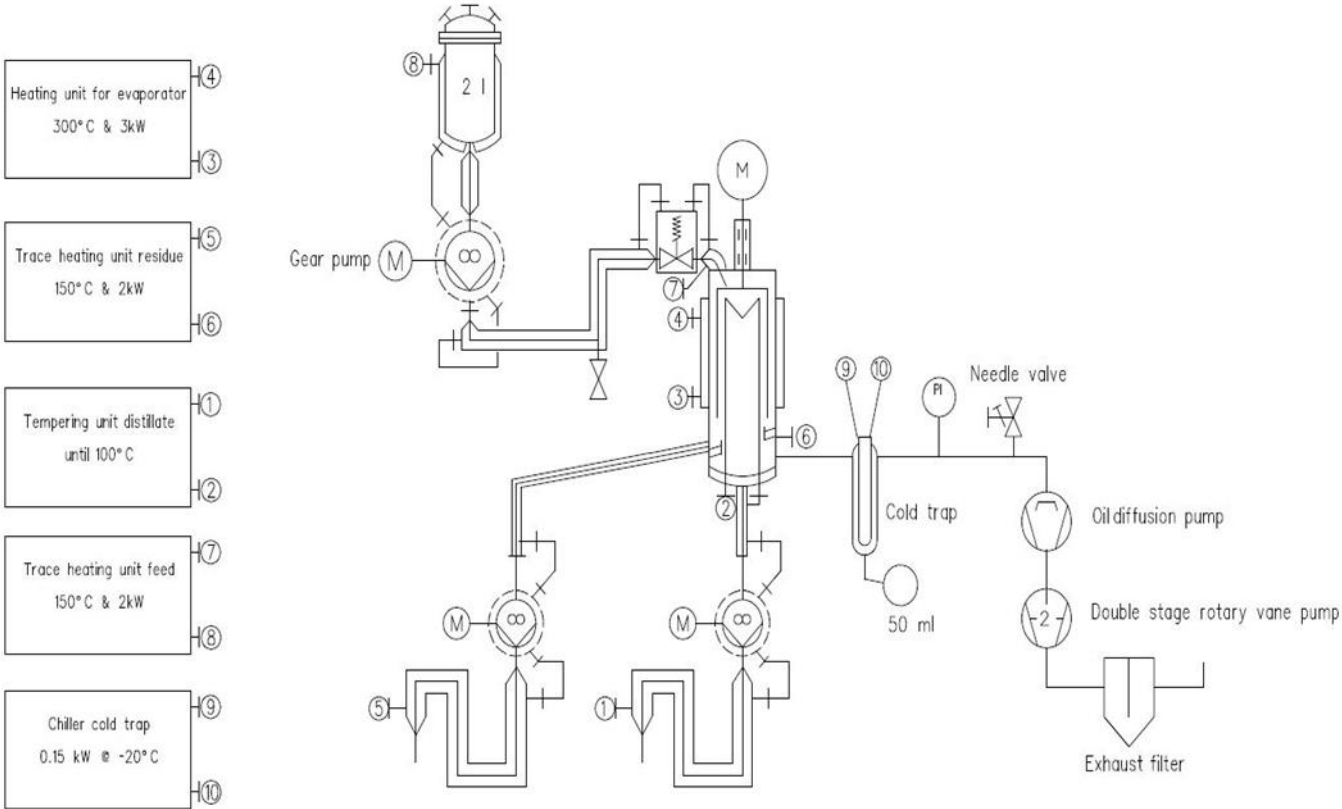
Figure A.5: Setup scheme of SPD plant VKL-70-5 (VTA GmbH, Niederwinkling)

Figure A.6: Layer thickness of palm oil on the the evaporator wall

Figure A.7: Solid condensat refined in a SPD plant at high (left) and low (right) evaporation temperatures
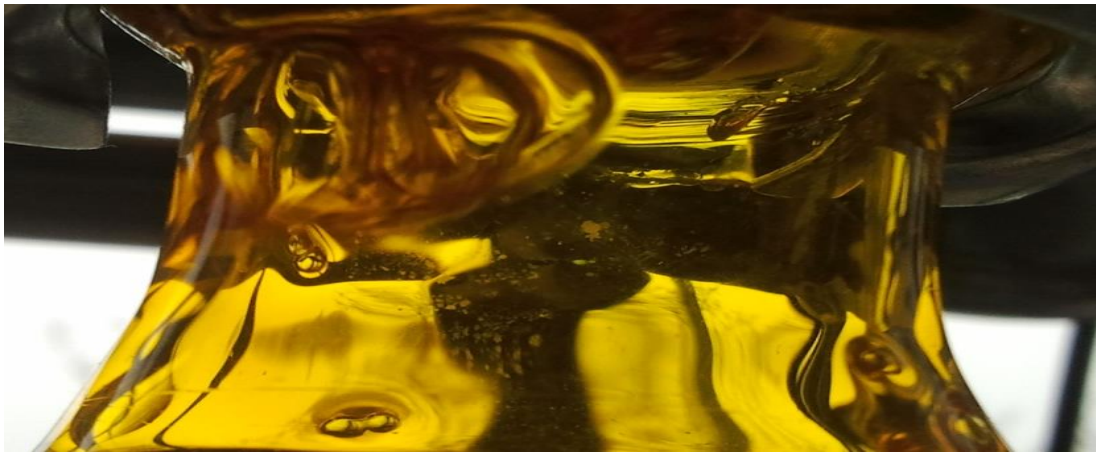
Figure A.8: Blistering within palm oil when entering an SPD plant with a high inflow speed

Annex B – Technical description of MATLAB functions

RSM literature example

Derivation of 1st grade model (implementation literature example):

A model of 1st grade is derived here by implementation of an example from the literature. The model is derived as exposed in chapter 2.3.2.1.3 by application of a full factorial design. The determination of parameter significances and model validation is achieved according to the methods applied in the literature example and as exposed in chapter 2.3.2.1.3.

Experimental_Design_Model_1st_Grade_Example_Literature.m :

This function calculates a full and fractional factorial design (normalized calculation matrix and non-normalized experimental setup) for the determination of a model of 1st grade. This file has to be run first for an RSM based on a model of 1st grade. RSM_Short_Path_Model_1st_Grade.m thus requires to run this file first (in RSM_Short_Path_Model_2nd_Grade.m the setup of the experimental design is already included).

- 3-6: The process variable range which is subject to analysis is defined.
- 8: The exponent m / m-q for the trial number calculation for full / fractional design is defined with m being the total number of process variables and q the fractional factor.
- 10-125: The setup of the trial and calculation matrix as well as the notation of effects referring to the columns of these matrices is achieved for both a full and fractional design (defined by the value of z).
- 19-23: The matrix Eff is introduced that contains the notation of the variable effects and interactions referring to the corresponding columns in the calculation matrix (compare matrix T_M below).
- 25-38: Definition of normalized process variable settings of the independent process variables for the trials of the experimental design. For the full factorial design these process variable settings are referred to as kernel for the design of next higher order. The kernel is part of the total calculation matrix T_M for the parameters of all model terms (single effects, interactions).
- 41-91: Calculation of the normalized settings of the interactions for the referring normalized process variable settings of the single trials (rows) in the matrix "Kernel". These settings for the interactions are saved to the matrix "I_A" and essential part of the calculation matrix from which the model parameters are calculated with respect to the experimental results for the experimental trials also defined in that calculation matrix (compare matrix "Kernel" as part of calculation matrix). Also the notations of the interactions referring to the columns in "I_A" are defined in Vector "eff".
- 96-106: Combination of dummy variable (for first term in the model equation), matrix "Kernel" (normalized settings of process variables) and matrix "I_A" (settings for the interactions) are combined yielding the calculation matrix "T_M". Saving calculation matrix and referring notations of effects to the workspace.
- 110-123: The non-normalized experimental settings are calculated and saved to the workspace.
- 129-140: Here the effects of the fractional design are defined that are set equal. The interactions of highest order of the 2^{m-q}-design are considered as being negligible and set equal to the additional variables of the corresponding 2^{m}-design.
- 143-184: Identification of the trials of the 2^{m}-design that form the corresponding 2^{m-q}-designs. Setup of these fractional designs which together form the 2^{m}-design.

RSM_Model_1st_Grade_Example_Literature.m:

Requires output from file "Experimental_Design_Model_1st_Grade_Example_Literature.m" which thus has to be run before.

- 5: Definition of whether a full and / or fractional factorial design has to be set up.
- 8-13: The critical values for the Bartlett-Criterion are defined. According to chapter 2.3.2.1.3 the Bartlett-Criterion is a criterion to check if the mean summed squares of the negligible model terms are a valid estimate for the error variance of experimentation. For validity the criterion has to be smaller than the

critical value. The Bartlett-Criterion and the referring critical value is displayed in the command window after each calculation cycle in order to check the procedure validity.

- 16-18: The critical values for the F-Test are defined. By application of that test for each model term that is considered as being essential the relative variance is calculated. If that value for the relative variance is larger than the referring critical value of the F-distribution then the parameter / model term is significant and remains in the equation (otherwise not). The relative variances as well as the critical F-values are displayed in the command window after each calculation cycle to check which parameters / model terms are significant.

- 20-21: Here the negligible (pos_1) and considered (pos_2) model parameters are defined. The negligible parameters are applied for the estimation of the error variance and calculation of the Bartlett-Criterion. At first the interactions of highest order / interactions that are considered to be negligible are chosen and the program is run. In the command window then the relative variance and critical values for the F-Test are displayed to identify significant parameters / model terms. The insignificant parameters with the lowest error variance are then incorporated into the negligible model terms in vector "pos_1" (also correct "pos_2" in this sense!) and the relative variances of the remaining model terms in "pos_2" are calculated again by running the program. This procedure is repeated until only significant model terms remain.

- 23-26: Definition of the experimental data for four trials referring to the order and definition of experiments in matrix "T_M".

- 31-38: Tr_1 contains the numbers of the experiments in T_M which are subject to further analysis dependent on whether a full or fractional design is applied (defined by "mode_design").

- 42-44: For each process setting calculation of the mean from all four experiment repetitions.

- 48-59: Calculation of the parameter estimates for all model terms of the model of 1st grade.

- 64-69: Calculation of the summed squares of the negligible model terms and the average of these summed squares. These values are required for calculation of the Bartlett-Criterion.

- 73-77: Calculation of the log summed squares of the negligible model terms. This value is required for calculation of the Bartlett-Criterion.

- 78-83: Calculation of Bartlett-Criterion. Display of Bartlett-Criterion, degrees of freedom and critical values in command window to check if the average of the summed squares of the negligible parameters (calculated in line 64-69) is a valid estimate of the error variance which then can be applied for calculation of the relative variances of the parameter terms.

- 87-91: Calculation of the summed squares of the considered / non-negligible model terms.

- 93-98: Calculation of the relative variances of the considered / non-negligible parameter terms which are applied for determination of model term significance via F-test.

- 100-107: Display of relative variances, degrees of freedom, and critical value for the F-test. For significant model terms / parameters the relative variance value is larger than the critical value.

- 109-118: Setup of parameter vector containing all parameter values for the significant model terms. For non-significant model terms the parameter value is set to zero. The order of parameters in "para_vec" follows the order and notations in vector "Order_and_Specifications_of_Effects_Full_Design_Large" which has been saved to the workspace before.

- 119: Calling function "Validation_Model_1st_Grade_Example_Literature.m" in which the model is further analyzed.

Validation_Model_1st_Grade_Example_Literature.m:

- 5-14: Calculation of the model estimates for the different process variable settings of the experimental trials in matrix "T_M" (experimental design applied). Calculation of the summed squares QS_L for model validation.

- 16-23: Calculation of the summed squares QS_F for model validation.

- 25: Calculation of the relative variance for model validation via F-test. The model is valid if the relative variance is smaller or equal to the critical F-value.

Derivation of 2nd grade model (implementation of literature example):

A model of 2nd grade is derived here by implementation of an example from the literature. The model is derived as exposed in chapter 2.3.2.1.4 by application of an orthogonal central composite design. The determination of parameter significances and model validation is achieved according to the methods applied in the literature example and as exposed in chapter 2.3.2.1.4.

RSM_Model_2nd_Grade_Example_Literature.m:

- 7-9: The critical values for the F-Test are defined. By application of that test for each model term that is considered as being essential the relative variance is calculated. If that value for the relative variance is larger than the referring critical value of the F-distribution then the parameter / model term is significant and remains in the equation (otherwise not). The relative variances as well as the critical F-values are displayed in the command window after each calculation cycle to check which parameters / model terms are significant.
- 11-12: Here the negligible (pos_1) and considered (pos_2) model parameters are defined. The negligible parameters are applied for the estimation of the error variance and calculation of the Bartlett-Criterion. At first the interactions of highest order / interactions that are considered to be negligible are chosen and the program is run. In the command window then the relative variance and critical values for the F-Test are displayed to identify significant parameters / model terms. The insignificant parameters with the lowest error variance are then incorporated into the negligible model terms in vector "pos_1" (also correct "pos_2" in this sense!) and the relative variances of the remaining model terms in "pos_2" are calculated again by running the program. This procedure is repeated until only significant model terms remain.
- 15-20: The critical values for the Bartlett-Criterion are defined. The Bartlett-Criterion is a criterion to check if the mean summed squares are a valid estimate for the variance of the error of experimentation. For validity the criterion M has to be smaller than the critical value. The Bartlett-Criterion and the referring critical value is displayed in the command window after each calculation cycle in order to check the procedure validity.
- 23: Experimental data according the to the order of experiments in matrix "T_M" (rows).
- 25-36: Definition of minimum, maximum and range to be investigated for the process variables.
- 38-43: Definition of the values for the characteristic constants of the orthogonal central composite design which are essential for further calculations and analysis such as calculation of parameter estimates.
- 49-62: Definition of normalized process variable settings of the independent process variables for the trials of the full factorial experimental design which is the kernel orthogonal central composite design applied here. The kernel is part of the total calculation matrix T_M for the parameters of all model terms (single effects, quadratic effects, interactions).
- 65-73: Definition of the normalized process variable settings for the star points.
- 76: Definition of the normalized process variable setting for the center point.
- 80: Combining normalized process variable settings for the kernel (full factorial design, star points and center point). The resulting matrix actually is the trial matrix that only contains the normalized settings of the process variables (the single linear effects respectively) for the experiments (rows).
- 84-93: Calculation of the normalized settings of the interactions for the referring normalized process variable settings of the single trials (rows). These settings for the interactions are saved to the matrix "I_A_1" and are essential part of the calculation matrix from which the model parameters are calculated with respect to the experimental results for the experimental trials also defined in that calculation matrix.
- 96-101: Calculation of the normalized settings of the squared effects for the referring normalized process variable settings of the single trials (rows). These settings for the interactions are saved to the matrix "C_E" and are essential part of the calculation matrix from which the model parameters are calculated with respect to the experimental results for the experimental trials also defined in that calculation matrix.
- 105: Setup of the calculation matrix "T_M" that shows for each experimental trial the settings of the process variables / model terms. Setup of "T_M" by combination of its parts (single effects, interactions, squared effects) that have been calculated above.
- 107-114: Calculation of the characteristic constants c and m for calculation of the parameter estimates.
- 118-147: Calculation and display (in workspace) of model parameters.
- 153-168: Calculation of the summed squares of the model terms.
- 170-175: Calculation of the average of the summed squares referring to the negligible model terms. This value is an estimate of the error variance if the Bartlett-Criterion is fulfilled.

- 177-181: Calculation of the log summed squares of the negligible model terms. This value is required for calculation of the Bartlett-Criterion.
- 182-187: Calculation of Bartlett-Criterion. Display of Bartlett-Criterion, degrees of freedom and critical values in command window to check if the average of the summed squares of the negligible parameters (calculated in line 170-175) is a valid estimate of the error variance which then can be applied for calculation of the relative variances of the parameter terms.
- 189-192: Calculation of the relative variances of the considered / non-negligible parameter terms which are applied for determination of model term significance via F-test.
- 193-198: Display of relative variances, degrees of freedom, and critical value for the F-test. For significant model terms / parameters the relative variance value is larger than the critical value.
- 138-147: Setup of parameter vector containing all parameter values for the significant model terms. For non-significant model terms the parameter value is set to zero.
- 216-254: In this section significant parameters are identified from the relative variances by application of an error variance estimate that is calculated with respect to the assumption that the model is adequate. This section has to be activated when this method is applied. The preceding method (estimation of error variance from negligible model terms) does not have to be inactivated as its results for error variance and relative variance of parameters are overwritten by following methods.
- 216-224: Calculation of model estimates / solutions for each experimental process variable setting by application of all model terms.
- 226-236: Calculation of the error variance estimate from the model estimates / solutions and the referring experimental observations.
- 233-236: Calculation of the relative variance for each model term with respect to the error variance estimate and the summed squares of the model terms.
- 238-241: Display of relative variance and critical value of an F-Test to determine significant parameters / model terms.
- 243-254: Sorting out significant parameters / model terms. Position and value of the parameter / model term are saved to the vectors "pos" and "para_vec".
- 259-295: In this section significant parameters are not identified from the relative variance by a F-Test but alternatively from the individual confidence intervals by a t-Test. The error variance estimate applied here is determined by one of the estimation methods implemented above (negligible model terms or assumption of model adequacy) dependent on whether the latter method is activated or not. This section has to be activated if this method is applied.
- 259-274: Calculation of the variances of the parameters / model terms.
- 275: Definition of the critical value for the t-Test.
- 277-282: Calculation of the individual confidence interval values for the parameters / model terms which are compared to the critical value as a criterion for decision on significance (via t-Test).
- 283-295: Sorting out significant parameters. Position and value of significant parameters are saved to the vectors "para_vec" and "pos".

Analysis of 2nd grade model:

Analysis_Model_2nd_Grade_Example_Script.m:

- 4-12: Definition of minimum and maximum of the normalized variable range which is subject to modeling and analysis via RSM. In non-transformed x-variables.
- 14-21: Definition of the corner points of the range of process variable settings (normalized) that is subject to modeling by RSM and the experimental design respectively. In this three-dimensional case these are the corner points of the cube representing the full factorial design for three process variables.
- 23-48: Plot of the border lines of the spectrum / cube of modeling and analysis.
- 48-59: Axes labeling, definition of title and legend.
- 62-64: Definition of the matrices and vectors in the matrix form of the model equation given by Eq. 5.4.
- 65: Calculation of the extremum variable setting according to Eq. 5.7.
- 66: Calculation of eigenvalue Matrix Λ and normalized eigenvector matrix M.
- 68-69: Calculation of function value from non-transformed coordinates at corner point P1 and extremum coordinates.
- 71-76: Plot of extremum in non-transformed coordinates / coordinate system.
- 78-84: Definition of the center points of the side planes of the cube representing the area of process variable settings that are subject to analysis / modeling.

- 86-89: Definition of the direction vectors of the axes of the non-transformed coordinate system.
- 93-154: For each z-axis the intersection with each side plane of the cube (range of analysis) is calculated. Also read in this context file / program description "test.m".
- 145-151: For each z-axis and side plane is checked if the intersection is located on the side plane of the cube.
- 157-170: Plot of intersections and z-axes into the figure in x-coordinates.
- 173-218: The direction of the z-axes in the x-coordinate system are checked and labeled in z-coordinates. In other words: Until now just the origin of the z-coordinate system and the intersections with the cube side planes are known and can be plot (transformed into x-coordinates) into the x-coordinate system / figure. It is not known so far in which direction of the z-axes which are plot into the x-coordinate system the value of z decreases and increases. This has to be checked by calculating the direction vector between origin of z-coordinate system and intersections in z-coordinates.
- 173-190: For each intersection of z-axis and cube side plane the direction vector in z-coordinates from intersection to the origin of the z-coordinate system is calculated to identify the axes directions.
- 192-218: To each intersection of z-axis and cube side plane the referring z-axis direction (algebraic sign of z-values on this side of the z-coordinate system origin) is plot into the figure.
- 220-245: For each intersection of a z-axis with a cube side plane the x- and z-coordinates as well as the referring function / target value is calculated. The latter is plot into the figure right next to the referring intersection. This is essential to identify how the target values depend on z-coordinate variation in a certain direction.
- 249-292: For further analysis the Z_2-Z_3-plane is plot with referring target values into the x-coordinate system. The point within that plane at which the target value reaches a maximum is calculated and plot with x-coordinates and function value into the figure.
- 296-361: The model is plot for a defined target value (either extremum (origin of z-coordinate system) or maximum within range of analysis can be chosen in line 337-338) in z-variables. The plot shows all process variable settings in z-coordinates at which the referring target value is reached.
- 296-329: For the range of normalized x-coordinates which is subject to analysis the corresponding range of z-coordinates is calculated. The latter is required to limit the analysis / plot of the canonic equation form to the range of practical relevance.
- 332-345: Calculation of the z-coordinates of the canonic function form for a defined target value (difference between target value and extremum respectively).
- 346-361: 3-D plot of the figure in the z-coordinate system.

test.m:

- In this file the matrix algebra method applied here to calculate the intersection between a plane and axis in a three-dimensional room is verified by a simple example in which the intersection coordinates are known before and have to be recovered from the parameters characterizing axis and plane (direction vectors, point coordinates). The calculation procedure is described sufficiently in the file. Thus no further explanations are required here.

RSM Short-Path-Distillation

The following program / algorithms achieve the derivation of linear stochastic process models of 1st and 2nd grade according to the theory exposed in chapter 2.3.2.1.3 and 2.3.2.1.4. The process models derived here are then subject to Bayesian model fit, discrimination, validation and analysis.

Experimental_Design_Model_1st_Grade.m :

This function calculates a full and fractional factorial design (normalized calculation matrix and non-normalized experimental setup) for the determination of a model of 1st grade. This file has to be run first for an RSM based on a model of 1st grade. RSM_Short_Path_Model_1st_Grade.m thus requires to run this file first (in RSM_Short_Path_Model_2nd_Grade.m the setup of the experimental design is already included).

- 3-7: The process variable range which is subject to analysis is defined.
- 9: The exponent m / m-q for the trial number calculation for full / fractional design is defined with m being the total number of process variables and q the fractional factor.

- 11-127: The setup of the trial and calculation matrix as well as the notation of effects referring to the columns of these matrices is achieved for both a full and fractional design (defined by the value of z).
- 20-25: The matrix Eff is introduced that contains the notation of the variable effects and interactions referring to the corresponding columns in the calculation matrix (compare matrix T_M below).
- 27-40: Definition of normalized process variable settings of the independent process variables for the trials of the experimental design. For the full factorial design these process variable settings are referred to as kernel for the design of next higher order. The kernel is part of the total calculation matrix T_M for the parameters of all model terms (single effects, interactions).
- 43-93: Calculation of the normalized settings of the interactions for the referring normalized process variable settings of the single trials (rows) in the matrix "kernel". These settings for the interactions are saved to the matrix "I_A" and are essential part of the calculation matrix from which the model parameters are calculated with respect to the experimental results for the experimental trials also defined in that calculation matrix (compare matrix "Kernel" as part of calculation matrix). Also the notations of the interactions referring to the columns in "I_A" are defined in Vector "eff".
- 98-108: Combination of dummy variable (for first term in the model equation), matrix "Kernel" (normalized settings of process variables) and matrix "I_A" (settings for the interactions) are combined yielding the calculation matrix "T_M". Saving calculation matrix and referring notations of effects to the workspace.
- 112-125: The non-normalized experimental settings are calculated and saved to the workspace.
- 131-142: Here the effects of the fractional design are defined that are set equal. The interactions of highest order of the 2^{m-q}-design are considered as being negligible and set equal to the additional variables of the corresponding 2^m-design.
- 145-186: Identification of the trials of the 2^m-design that form the corresponding 2^{m-q}-designs. Setup of these fractional designs which together form the 2^m-design.

RSM_Short_Path_Model_1st_Grade.m:

Requires output from file "Experimental_Design_Model_1st_Grade.m" which therefore has to be run before.

- 5: Definition of either a full factorial or fractional factorial design is applied.
- 7-22: Definition of minimum, maximum and range to be investigated for the process variables.
- 25-30: The critical values for the Bartlett-Criterion are defined. The Bartlett-Criterion is a criterion to check if the mean summed squares are a valid estimate for the variance of the error of experimentation. For validity the criterion M has to be smaller than the critical value. The Bartlett-Criterion and the referring critical value is displayed in the command window after each calculation cycle in order to check the procedure validity.
- 33-35: The critical values for the F-Test are defined. By application of that test for each model term that is considered as being essential the relative variance is calculated. If that value for the relative variance is larger than the referring critical value of the F-distribution then the parameter / model term is significant and remains in the equation (otherwise not). The relative variances as well as the critical F-values are displayed in the command window after each calculation cycle to check which parameters / model terms are significant.
- 37-42: Here for each modeling approach (acid value, rancimat, peroxide value and tocopherols) the negligible (pos_1) and considered (pos_2) model parameters are defined. The negligible parameters are applied for the estimation of the error variance and calculation of the Bartlett-Criterion. At first the interactions of highest order / interactions that are considered to be negligible are chosen and the program is run. In the command window then the relative variance and critical values for the F-Test are displayed to identify significant parameters / model terms. The insignificant parameters with the lowest error variance are then incorporated into the negligible model terms in vector "pos_1" (also correct "pos_2" in this sense!) and the relative variances of the remaining model terms in "pos_2" are calculated again by running the program. This procedure is repeated until only significant model terms remain.
- 44-51: Definition of the experimental data for two trials referring to the order and definition of experiments in matrix "T_M".
- 53-59: Tr_1 contains the numbers of the experiments in T_M which are subject to further analysis dependent on whether a full or fractional design is applied (defined by "mode_design").
- 61-63: Definition of experimental values that are deleted due to a reasonable outlier.
- 77-89: Calculation of the parameter estimates for all model terms of the model of 1st grade.

- 93-98: Calculation of the summed squares of the negligible model terms and the average of these summed squares. These values are required for calculation of the Bartlett-Criterion.
- 102-106: Calculation of the log summed squares of the negligible model terms. This value is required for calculation of the Bartlett-Criterion.
- 107-112: Calculation of Bartlett-Criterion. Display of Bartlett-Criterion, degrees of freedom and critical values in command window to check if the average of the summed squares of the negligible parameters (calculated in line 93-98) is a valid estimate of the error variance which then can be applied for calculation of the relative variances of the parameter terms.
- 116-120: Calculation of the summed squares of the considered / non-negligible model terms.
- 122-127: Calculation of the relative variances of the considered / non-negligible parameter terms which are applied for determination of model term significance via F-test.
- 128-136: Display of relative variances, degrees of freedom, and critical value for the F-test. For significant model terms / parameters the relative variance value is larger than the critical value.
- 138-147: Setup of parameter vector containing all parameter values for the significant model terms. For non-significant model terms the parameter value is set to zero. The order of parameters in "para_vec" follows the order and notations in vector "Order_and_Specifications_of_Effects_Full_Design_Large" which has been saved to the workspace before.
- 148: Calling function "Analysis_Model_1st_Grade_Short_Path.m" in which the model is further analyzed.

Analysis_Model_1st_Grade_Short_Path.m:

- 5-25: Calculation of the model estimates for the different process variable settings of the experimental trials in matrix "T_M" (experimental design applied). Calculation of the summed squares QS_L for model validation. This is achieved with respect to experimental data that has been deleted as outlier.
- 27-38: Calculation of the summed squares QS_F for model validation. This is achieved with respect to experimental data that has been deleted as outlier.
- 40: Calculation of the relative variance for model validation via F-test. The model is valid if the relative variance is smaller or equal to the critical F-value.
- 42-45: Definition of process variable spectrum and interval at which the function value is calculated for the three-dimensional function plot.
- 46: In case of more than two process variables in the function that is subject to the three-dimensional plot only two process variables stay variable while the others have to be fixed to a defined value.
- 47: For the process variables that are subject to the three-dimensional plot a two-dimensional grid is defined from the corresponding variable spectrum / interval (line 42-45). Each intersection of the grid is a point for which the coordinates vor the two process variables are saved. The grid is equal to the plane in the three-dimensional plot that is spanned by the coordinate axis of the referring process variables. Ensure to define the input and output in function "meshgrid" for the right process variables.
- 48-56: Calculation of the function value for each grid point / process variable coordinates. Ensure to activate the right function to be analyzed here and that the right process variables in the function stay variable while the others are fixed to a defined value.
- 58-72: Three dimensional plot of the function. Ensure to activate the corresponding axes labeling and title here.
- 75-77: Searching for the process variable setting at which the function value is minimized. To search for a maximum the function equation has to be multiplied by -1 in "Func_1st_Grade_Short_Path.m".

Func_1st_Grade_Short_Path.m:

- 3-9: The minimization routine "fminsearch" also proposes automatically process variable settings which are outside the normalized range of process settings that is subject to the design of experiments and process analysis. As the experimental data and all results derived from that are only valid within that variable range the extremum within that range has to be identified. Thus all proposals for the variable setting that are outside that range have to be limited to the corresponding limiting values.
- 11-16: Calculation of the function values for the proposed process variable setting. Saving the optimum values to the workspace. Ensure to activate the right model with the right algebraic notation here depending on whether a minimum or maximum shall be identified. Be aware that vector X contains the process variables of the model in numeric order. Therefore it might for example happen that in third place of vector X the setting for the forth process variable "pump power" is saved if the third process variable "stirrer rotation" is not significant with respect to the target / function value.

RSM_Short_Path_Model_2nd_Grade.m:

- 7-9: The critical values for the F-Test are defined. By application of that test for each model term that is considered as being essential the relative variance is calculated. If that value for the relative variance is larger than the referring critical value of the F-distribution then the parameter / model term is significant and remains in the equation (otherwise not). The relative variances as well as the critical F-values are displayed in the command window after each calculation cycle to check which parameters / model terms are significant.

- 11-16: Here the negligible (pos_1) and considered (pos_2) model parameters for the different modeling approaches (rancimat, tocopherol, acid value) are defined. The negligible parameters are applied for the estimation of the error variance and calculation of the Bartlett-Criterion. At first the interactions of highest order / interactions that are considered to be negligible are chosen and the program is run. In the command window then the relative variance and critical values for the F-Test are displayed to identify significant parameters / model terms. The insignificant parameters with the lowest error variance are then incorporated into the negligible model terms in vector "pos_1" (also correct "pos_2" in this sense!) and the relative variances of the remaining model terms in "pos_2" are calculated again by running the program. This procedure is repeated until only significant model terms remain.

- 19-24: The critical values for the Bartlett-Criterion are defined. The Bartlett-Criterion is a criterion to check if the mean summed squares are a valid estimate for the variance of the error of experimentation. For validity the criterion M has to be smaller than the critical value. The Bartlett-Criterion and the referring critical value is displayed in the command window after each calculation cycle in order to check the procedure validity.

- 27-32: Experimental data for the different target values / modeling approaches according to the order of experiments in matrix "T_M".

- 37-38: Definition of experimental values that are sorted out as reasonable outliers.

- 40-50: Calculation of the vector "Y" that contains for each experimental setting the mean of all repetitions of that experiment with respect to reasonable outliers.

- 53-68: Definition of minimum, maximum and range to be investigated for the process variables.

- 70-75: Definition of the values for the characteristic constants of the orthogonal central composite design that are essential for further calculations and analysis such as calculation of parameter estimates.

- 81-94: Definition of normalized process variable settings of the independent process variables for the trials of the full factorial experimental design which is the kernel orthogonal central composite design applied here. The kernel is part of the total calculation matrix T_M for the parameters of all model terms (single effects, quadratic effects, interactions).

- 97-105: Definition of the normalized process variable settings for the star points.

- 108: Definition of the normalized process variable setting for the center point.

- 112: Combining normalized process variable settings for the kernel (full factorial design, star points and center point). The resulting matrix actually is the trial matrix that only contains the normalized settings of the process variables (the single linear effects respectively) for the experiments (rows).

- 115-124: Calculation of the normalized settings of the interactions for the referring normalized process variable settings of the single trials (rows). These settings for the interactions are saved to the matrix "I_A_1" and are essential part of the calculation matrix from which the model parameters are calculated with respect to the experimental results for the experimental trials also defined in that calculation matrix.

- 127-132: Calculation of the normalized settings of the squared effects for the referring normalized process variable settings of the single trials (rows). These settings for the interactions are saved to the matrix "C_E" and are essential part of the calculation matrix from which the model parameters are calculated with respect to the experimental results for the experimental trials also defined in that calculation matrix.

- 136: Setup of the calculation matrix "T_M" that shows for each experimental trial the settings of the process variables / model terms. Setup of "T_M" by combination of its parts (single effects, interactions, squared effects) that have been calculated above.

- 138-145: Calculation of the characteristic constants c and m for calculation of the parameter estimates.

- 149-178: Calculation and display (in workspace) of model parameters.

- 184-199: Calculation of the summed squares of the model terms.

- 201-206: Calculation of the average of the summed squares referring to the negligible model terms. This value is an estimate of the error variance if the Bartlett-Criterion is fulfilled.

- 208-212: Calculation of the log summed squares of the negligible model terms. This value is required for calculation of the Bartlett-Criterion.
- 213-218: Calculation of Bartlett-Criterion. Display of Bartlett-Criterion, degrees of freedom and critical values in command window to check if the average of the summed squares of the negligible parameters (calculated in line 170-175) is a valid estimate of the error variance which then can be applied for calculation of the relative variances of the parameter terms.
- 220-223: Calculation of the relative variances of the considered / non-negligible parameter terms which are applied for determination of model term significance via F-test.
- 224-229: Display of relative variances, degrees of freedom, and critical value for the F-test. For significant model terms / parameters the relative variance value is larger than the critical value.
- 231-243: Setup of parameter vector containing all parameter values for the significant model terms. For non-significant model terms the parameter value is set to zero.
- 247-285: In this section significant parameters are identified from the relative variances by application of an error variance estimate that is calculated with respect to the assumption that the model is adequate. This section has to be activated when this method is applied. The preceding method (estimation of error variance from negligible model terms) does not have to be inactivated as its results for error variance and relative variance of parameters are overwritten by following methods.
- 247-255: Calculation of model estimates / solutions for each experimental process variable setting by application of all model terms.
- 257-262: Calculation of the error variance estimate from the model estimates / solutions and the referring experimental observations.
- 264-267: Calculation of the relative variance for each model term with respect to the error variance estimate and the summed squares of the model terms.
- 270-272: Display of relative variance and critical value of an F-Test to determine significant parameters / model terms.
- 274-285: Sorting out significant parameters / model terms. Position and value of the parameter / model term are saved to the vectors "pos" and "para_vec".
- 289-321: In this section significant parameters are identified from the relative variances by application of an error variance estimate that is calculated from multiple realizations of an experiment. This section has to be activated when this method is applied. The preceding methods (estimation of error variance from negligible model terms and estimation of error variance with the assumption that the model is true) does not have to be inactivated as its results for error variance and relative variance of parameters are overwritten by following methods.
- 289-300: Calculation of the error variance from multiple realizations of the experimental setting.
- 302-305: Calculation of the relative variance of the model terms / parameters.
- 306-309: Display of relative variances and critical value. Saving relative variances to the workspace.
- 310-321: Sorting out significant parameters. Position and value of significant parameters are saved to the vectors "para_vec" and "pos".
- 326-362: In this section significant parameters are not identified from the relative variance by a F-Test but alternatively from the individual confidence intervals by a t-Test.). The error variance estimate applied here is determined by one of the estimation methods implemented above (negligible model terms or assumption of model adequacy) dependent on whether the latter method is activated or not. This section has to be activated if this method is applied.
- 326-340: Calculation of the variances of the parameters / model terms.
- 242: Definition of the critical value for the t-Test.
- 344-348: Calculation of the individual confidence interval values for the parameters / model terms which are compared to the critical value as a criterion for decision on significance (via t-Test).
- 350-362: Sorting out significant parameters. Position and value of significant parameters are saved to the vectors "para_vec" and "pos".
- 367: Calling function Analysis_Model_2nd_Grade_Short_Path.m for further model analysis.

Analysis_Model_2nd_Grade_Short_Path.m:

- 5-25: Calculation of the model estimates for the different process variable settings of the experimental trials in matrix "T_M" (experimental design applied). Calculation of the summed squares QS_L for model validation. This is achieved with respect to experimental data that has been deleted as outlier.
- 27-38: Calculation of the summed squares QS_F for model validation. This is achieved with respect to experimental data that has been deleted as outlier.

- 40: Calculation of the relative variance for model validation via F-test. The model is valid if the relative variance is smaller or equal to the critical F-value.
- 43-47: Searching for the process variable setting at which the function value is minimized. To search for a maximum the function equation has to be multiplied by -1 in "Func_2nd_Grade_Short_Path.m".
- 50-53: Definition of process variable spectrum and interval at which the function value is calculated for the three-dimensional function plot.
- 54-56: In case of more than two process variables in the function that is subject to the three-dimensional plot only two process variables stay variable while the others have to be fixed to a defined value.
- 57: For the process variables that are subject to the three-dimensional plot a two-dimensional grid is defined from the corresponding variable spectrum / interval (line 50-53). Each intersection of the grid is a point for which the coordinates vor the two process variables are saved. The grid is equal to the plane in the three-dimensional plot that is spanned by the coordinate axis of the referring process variables. Ensure to define the input and output in function "meshgrid" for the right process variables.
- 58-67: Calculation of the function value for each grid point / process variable coordinates. Ensure to activate the right function to be analyzed here and that the right process variables in the function stay variable while the others are fixed to a defined value.
- 69-83: Three dimensional plot of the function. Ensure to activate the corresponding axes labeling and title here.

Func_2nd_Grade_Short_Path.m:

- 5-15: The minimization routine "fminsearch" also proposes automatically process variable settings which are outside the normalized range of process settings that is subject to the design of experiments and process analysis. As the experimental data and all results derived from that are only valid within that variable range the extremum within that range has to be identified. Thus all proposals for the variable setting that are outside that range have to be limited to the corresponding limiting values.
- 17-23: Calculation of the function values for the proposed process variable setting. Saving the optimum values to the workspace. Ensure to activate the right model with the right algebraic notation here depending on whether a minimum or maximum shall be identified. Be aware that vector X contains the process variables of the model in numeric order. Therefore it might for example happen that in third place of vector X the setting for the forth process variable "pump power" is saved if the third process variable "stirrer rotation" is not significant with respect to the target / function value.

Bayesian fit, discrimination and analysis of growth models

In the following the m-files of the program are described in the order they are applied to the computational process. The numbers indicate the line of the specific m-file currently explained.

Mltpl_Chn_Rndm_Wlk_Cov_Updt_Growth_Model.m:

With this file the whole computational process of the Bayesian method applied here is started. This file provides estimates of the expected value (vector "mean_vec") as well as the variance (vector "var_vec") of the model parameter probability distribution at defined steps during the convergence of this distribution towards the posterior. These estimates are needed to graphically expose the convergence process and to identify the point of chain convergence. This file thus also provides a representative sample of parameter vectors from the posterior distribution which is taken from the parallel Markov-Chains after convergence (vector "Par_Smpl"). Additionally a vector "lik" is provided that contains the likelihood values of the parameter vectors contained in the vector "Par_Smpl". The latter two vectors are applied to the calculation of expected value and variance of the posterior model parameter distribution as well as the relative model probability for model discrimination. The output of this file which is basically the output of the Bayesian Markov-Chain process is saved to the workspace. It is necessary for further data analysis by Bayesian means in the file "Data_Analysis.m" which loads the data from the workspace automatically. It is recommended to save the output of this file on a data storage medium. In this case a Bayesian model fit, analysis and discrimination by file "Data_Analysis.m" can be achieved without running the whole Markov-Chain process again which is extremely time consuming.

- 4-8: Calls file "Input_Growth_Model.m" to introduce the basic data input necessary for the further computational process. In the following this data is defined as data matrix "input" which is assigned to the workspace and is globally valid / accessible to all associated m-files.
- 10-13: Here the process variables of the Bayesian Markov-Chain-Monte-Carlo process as well as its specific acceleration method are set up. These variables are the number of updates / re-estimates of the covariance matrix of the proposal density, the number of tunes of the covariance matrix of the proposal density within each re-estimation cycle of that matrix, the number of steps of the Markov-Chain within each tuning cycle of that matrix and finally the number of parallel Markov-Chains.
- 15: In the following loop the variable "t" defines the model treated. For model discrimination both models (t=1:2) have to be analyzed by Bayesian Markov-Chain-Monte-Carlo yielding the output mentioned above for each model.
- 17-21: Vector "Param" contains the start values for the parallel Markov-Chains. In the following the vector "Cov_Mat" takes up the re-estimated covariance matrix of the proposal density which is updated in each cycle of the loop initiated in line 23. The vectors "mean_vec" and "var_vec" take up the estimates of expected value und variance of the converging model parameter distribution as exposed above. The vector "par_tot" takes up the parameter vectors for each step of all parallel Markov-Chains. Except the last step of the preceding cycle this vector is cleared after each covariance re-estimation cycle to safe CPU and memory. In the last cycle (for which chain convergence has to be assured) the content of vector "par_tot" is saved as "Par_Smpl" (line 71) which is needed for further computation as exposed above.
- 23: This loop starts 1000 parallel Markov-Chains with 20 update cycles of the covariance matrix of the proposal density. Each update cycle contains of 10 tuning cycles of that matrix and each tuning cycle of 20 Markov-Chain steps. Thus each Markov-Chain consists of 10.000 steps in total. Covariance matrix update and tuning is implemented here according to the principle exposed in chapter.
- 29: Here m-file "Mltpl_Chn_Rndm_Wlk_Growth_Model.m" is called which achieves 10 tuning steps of the covariance matrix of the proposal density within each update cycle. Within in each tuning cycle of file "Mltpl_Chn_Rndm_Wlk_Growth_Model.m" 20 steps are achieved for each of the parallel Markov-Chains. File "Mltpl_Chn_Rndm_Wlk_Growth_Model.m" provides the vector "par_tot" (see above) and the vector "lik" that contains the referring likelihood values for the parameter vectors in "par_tot".
- 32-59: m-file "Mltpl_Chn_Rndm_Wlk_Growth_Model.m" provides the output of 10 tuning cycles each consisting of 20 steps for each of the 1000 parallel Markov-Chains. The last 20 steps for each of the 1000 parallel Markov-Chains are used to estimate the expected value and covariance which represent an estimation of the current state of the parameter probability distribution converging to the posterior distribution. After each re-estimation cycle these distribution parameters are saved in the vectors 'mean_vec' and 'var_vec' for further use (see above).
- 61-63: The actual variance of the parameter probability distribution is applied to the covariance matrix of the proposal density for the next re-estimation cycle.
- 67-72: The vector 'par_tot' that contains the parameter vectors of all parallel Markov-Chains of each entire re-estimation cycle is cleared (saving CPU) except the last vector of each chain which is required as starting point for the next re-estimation cycle. For the last re-estimation cycle (completed convergence has to be assured by graphical analysis) the parameter vectors for all parallel Markov-Chains are saved in the vector "Par_Smpl" which is at completed convergence a representative sample of the required posterior parameter probability distribution.
- 76-156: Plot of the convergence of the parameter probability distribution towards the posterior distribution. Plots are achieved for both models and each of the characteristic distribution parameters (expected value and variance).

Mltpl_Chn_Rndm_Wlk_Growth_Model.m:

- For each re-estimation cycle m-file "Mltpl_Chn_Rndm_Wlk_Growth_Model.m" achieves 10 tuning steps of the covariance matrix of the proposal density with each step consisting of 20 steps each for 1000 parallel Markov-Chains. The procedure of running parallel Markov-Chains with iterative tuning of the proposal density is implemented here. File "Mltpl_Chn_Rndm_Wlk_Growth_Model.m" provides the vectors "par_tot" and "lik" whose further function has been already exposed above.
- 5: This loop achieves 10 tuning steps of the covariance matrix of the proposal density each being based on 20 steps of 1000 parallel Markov-Chains each.
- 10: File "Rndm_Wlk_Growth_Model.m" achieves 20 steps of 10000 parallel Markov-Chains each. This file provides for each tuning cycle the vectors "par", "Acc_Prob_Vec" and "Lik" which contain for each Markov-Chain the parameter vectors and the referring acceptance probability and likelihood value.
- 12-23: After each tuning cycle achieved by file "Rndm_Wlk_Growth_Model.m" the latest chain steps achieved in that file are added to vector "par_tot" for each of the parallel Markov-Chains.
- 25: In the context of line 12-23 the referring likelihood values of the parameter vectors, chain steps respectively, are added analogously.
- 27-40: Calculation of the average acceptance probability of the additional chain steps generated in the latest tuning cycle by file "Rndm_Wlk_Growth_Model.m". Tuning of the covariance matrix of the proposal density (average acceptance probability as decision criterion).

Rndm_Wlk_Growth_Model.m:

- For each tuning cycle of file "Mltpl_Chn_Rndm_Wlk_Growth_Model.m" file "Rndm_Wlk_Growth_Model.m" achieves 20 steps for 1000 parallel Markov-Chains. As output this file provides the vectors "par", "Acc_Prob_Vec" und "Lik" whose function and further processing has been exposed above.
- 6-9: The input data required for the Bayesian Markov-Chain process is uploaded.
- 10-12: The vectors taking up the process output subsequently are defined.
- 14-201: This loop achieves 1000 Markov-Chains one after another. Within that loop 20 chain steps are achieved.
- 16-21: Definition of the start vector for each chain. If the Markov-Chain has to be started (first re-estimation cycle and first tuning cycle) the input start vector is applied. Otherwise the last vector of the chain to be elongated (last vector of the last cycle) is applied as start vector. As all chains have to be subject to identical conditions, all chains are started (first re-estimation cycle and first tuning cycle) from the same start vector defined in the file "input.m".
- 25-197: This inner loop achieves 20 chain steps for each chain called up by the outer loop (line 14-201, see above).
- 27-89: In this statement for each chain step the latest accepted state / parameter vector of the chain actually treated is defined as well as the proposed new state / parameter vector of that chain.
- 27-58: For the first step of a chain (N=1) the latest accepted state is the start vector defined in line 16-21. Regarding the proposal of a new parameter vector as potential new chain state two situations have to be distinguished for the first chain step (N=1): The first situation (line 29-38) is that the program is still in the first covariance estimation cycle (a=1). Here no covariance of the actual state of the converging parameter probability distribution has been estimated so far as this is done at the end of each estimation cycle (see Mltpl_Chn_Rndm_Wlk_Cov_Updt_Growth_Model.m line 32-59 and 61-63). In this case the proposed candidate vector is initially sampled randomly from a Gaussian distribution with the start vectors defined in file "input.m" as mean and a standard deviation of 50%. This procedure is chosen to initially provide a broad parameter range for the Bayesian Markov-Chain process within which the random process can narrow the investigated parameter range to a coarse estimate of the range of values that are most possible in the light of the experimental data. The second situation (line 38-56) is that the program is in the second or an upper re-estimation cycle and thus a covariance matrix of the actual state of the converging parameter probability distribution has been already estimated (see Mltpl_Chn_Rndm_Wlk_Cov_Updt_Growth_Model.m line 32-59 and 61-63). The proposed parameter vector is in this case generated by random sampling from a Gaussian distribution with the latest estimated covariance matrix as variance and the start vector (compare line 16-21) as mean.
- 58-89: For the second and all following chain steps (N>1) two situations can be distinguished: The first situation (line 60-69) is that the program is still in the first covariance estimation cycle (a=1). For this case the same procedure for proposal of a candidate parameter vector is valid as exposed for the first chain step (N=1) above. The second situation (line 69-87) is that the program is in the second or an upper covariance re-estimation cycle. In this case the latest accepted state / parameter vector from a preceding chain step is defined as origin for the proposal of a candidate parameter vector to be accepted or rejected as next chain state. The proposed candidate is derived by random sampling from a Gaussian distribution with the latest accepted chain state as mean and the latest estimate of the covariance matrix as variance.
- 97-169: Calculation of prior probability and likelihood of the actual state and the parameter vector that has been proposed as next candidate. As prior knowledge about model parameterization a Gaussian distribution with the start vectors defined in file "input.m" as mean and a standard deviation of 25% is applied. The choice of such a rather small prior knowledge (instead the results of a residual least squares fit could be applied as improved prior knowledge about the mean combined with a standard deviation that has been calculated from a repeated residual least squares fit of the model to data that has been generated by repeated random addition of noise to the experimental data) has been made to improve the exposition of the performance of the Bayesian method applied here, especially the convergence towards the posterior distribution. All probabilities are calculated here as log values for the reasons exposed in the beginning of the description of file "Posterior_Ordinate_Growth_Model.m" below.
- 171-195: Implementation of the criterion and decision of candidate vector acceptance or rejection. The formula for the calculation of the criterion is converted here to calculate a log value for the criterion from the log probabilities. That log value is re-converted into a non-log value that is then applied for the acceptance-rejection-decision.
- 199: Saving the chain of accepted chain states / parameter vectors for the chain that has been actually treated by the loop from line 14-201.

Input_Growth_Model.m :

This file contains all input data on the models to be analyzed as well as the experimental data required for the computational process.

- 4-6: Experimental data of the microbial growth process. Substrate and cell concentration at the different time points of sampling.
- 7-10: Values of the model parameterization which are used as start values for the Markov-Chain process applied to the Bayesian determination of the most probable parameterization. These start values represent the prior knowledge on the process
- 12: Error / standard deviation of experimental analysis / measurement in percent of total value.
- 17-33: In order to show that the Bayesian method applied here is really able to identify the model that is more likely to describe the real underlying process it is assumed that the real underlying process is of Michaelis-Menten- or Moser-type with parameterization given in vector P_MM or P_Moser. With respect to the standard deviation of measurement given above probable experimental data of the process is generated. On the basis of that experimental data the Bayesian method should identify the Michaelis-Menten- or Moser-model as more probable compared to the other one. If the real experimental date in line 4-6 or the direct model solution (line 14-15) shall be subject to further computation line 17-33 have to be deactivated.
- 14-15: Calculation of the solution for cell and substrate concentration at the time points given above for a Michaelis-Menten- or Moser-model with the parameterization in vector P_MM or P_Moser. This data is considered as being representative for the real underlying process.
- 20-30: Generation of artificial experimental data by random sampling from a Gaussian probability distribution with the standard deviation given above. Noise is added randomly to the model solution calculated in line 14-15 to simulate a probable experimental outcome.
- 32-35: The experimental data applied to the computational process is re-defined if the direct model solution or the model solution with randomly added noise is applied as artificial experimental data. Activate / deactivate the corresponding lines.
- 39-50: Definition of prior knowledge on model parameterization.
- 39-46: RLS-fit of models to original or simulated experimental data. RLS-fit parameterization can be applied as prior knowledge on model parameterization. Inactivate when different prior knowledge is applied (compare line 47-50).
- 47-50: Activate / inactivate if either parameter start values or RLS-fit to real experimental data are applied as prior knowledge on model parameterization.

Start_ODE_Solver_MM.m:

- 5-15: Calculates the solution at the time points given by the vector tspan for a Michealis-Menten model with the parameterization given by the vector p.
- 18-25: In the loop all zero-concentrations are set to one as otherwise just zero-variances can be calculated yielding covariance matrices with a zero in the main diagonal with which no Gaussian probabilities can be calculated anymore causing an abort of the computational process. This approach is acceptable as such low concentrations are negligible compared to the concentrations and process regions of interest.

Start_ODE_Solver_Moser.m:

Analogous to description for file "Start_ODE_Solver_MM.m ".

MM_Model.m:

- Contains the Michaelis-Menten model which is called by the ODE-Solver to calculate the deviation of the cell and substrate concentration within the next step of the time interval.

Moser_Model.m:

- Contains the Moser model which is called by the ODE-Solver to calculate the deviation of the cell and substrate concentration within the next step of the time interval.

Data_Analysis.m

This file achieves the Bayesian model fit, analysis and discrimination based on the output of file "Mltpl_Chn_Rndm_Wlk_Cov_Updt_Growth_Model.m" which is the representative parameter sample from the posterior distribution and associated likelihood values as well as mean and variance values from the complete convergence process of the Markov-Chain. The data to be processed has to be provided directly by running file "Mltpl_Chn_Rndm_Wlk_Cov_Updt_Growth_Model.m" or has to be loaded into the workspace from a data storage medium on which the data output from a preceding run of that file has been saved before.

- 5-211: Estimation of the log marginal likelihoods for both models. The computations described in the following are thus just achieved for the model actually analyzed which is indicated by the value of variable t.
- 7-13: The input data is loaded upon which the Bayesian analysis achieved here and the applied Markov-Chain output is based (see "input.m" above).
- 17-23: The required input data and output of the Markov-Chains referring to the Michaelis-Menten model is loaded from the workspace.
- 27-33: The required input data and output of the Markov-Chains referring to the Moser model is loaded from the workspace.
- 37-43: An average of the posterior expected value and variance estimates is calculated as final estimate. These average values are saved to the workspace. The parameter G defines the number of estimate values to be taken for averaging. It is important to only apply estimates calculated from the converged chains. For that purpose the point of convergence has to be identified before graphically.
- 45-65: For all parameter vectors of the posterior sample a comparable posterior probability measure is calculated from the prior probability and likelihood. The denominator of posterior probability estimation (compare Eq. 2.18) is a normative constant which is equal for all parameter vectors of the posterior sample and, thus, not required to estimate the mode. From these posterior probability measures the posterior mode estimate is calculated which is further applied as Bayesian point estimate for model parameterization and as candidate vector for relative model probability estimation.
- 67-73: Posterior mode and variance estimates are saved to the workspace.
- 75-147: Plot of the convergence of the parameter probability distribution towards the posterior distribution. The convergence plots are achieved for expected value and variance estimates as distribution parameters and for both models.
- 149-169: For both models the model solutions for a Bayesian parameterization (most probable parameterization as calculated above) and additionally for a residual least squares fit of the model to the experimental data are calculated. File "Fit_Model_Equation_RLS.m" achieves the residual least squares fit of either model to the experimental data dependent on the actual setting of variable t that indicates which model is actually analyzed.
- 171-197: For the model actually analyzed (dependent on the actual value of variable "t") the model solution for a Bayesian parameterization and a residual least squares fit are plot together with the experimental data. This is done to graphically evaluate the goodness of fit and the differences in model solution between Bayesian and residual least squares parameterization.
- 199-217: Calculation of the relative model probability as model discrimination criterion by the method proposed by Chib and Jeliazkov.
- 199-200: The variance of the posterior parameter probability is defined and applied for the estimation of the marginal likelihood for each model according to the method proposed by Chib and Jeliazkov. According to that proposed method also the mode estimate of the posterior probability distribution is defined and applied as candidate vector for the estimation of the marginal likelihood as it exhibits the highest probability under that distribution. The choice of a candidate vector with a high probability under the posterior distribution is recommended.
- 203: Call of file "Posterior_Ordinate_Growth_Model.m". This file achieves the calculation of the posterior ordinate as well as the log prior probability and log likelihood of the candidate vector. To ensure a high precision of the marginal likelihood estimate a parameter sample from the posterior distribution has to be applied that is sufficiently representative. Thus, the whole representative sample taken from the Markov-Chains is applied for the posterior ordinate estimation. This comprises a sample of 200 parameter vectors from 1000 parallel Markov-Chains each which is by far sufficient even for a four dimensional probability distribution in case of the Moser-model. As far as possible by means of mathematical calculation rules, all computational procedures in the course of marginal likelihood estimation are achieved by use of log probabilities. This approach is necessary, as especially in case of likelihood calculations extremely low probabilities can occur which are not computational manageable anymore by the program.
- 203: Calculation of the log marginal likelihood from the output of "Posterior_Ordinate_Growth_Model.m".
- 205-209: Saving the log marginal likelihoods of the models to the workspace.
- 211: End of log marginal likelihood estimation loop.
- 213-214: Loading of the log marginal likelihood estimates for final calculation of the relative model probability.
- 216-219: Calculation of the relative model probability and saving to the workspace.

Fit_Model_Equation_RLS.m:

Achieves the residual least squares fit of either model to the experimental data dependent on the actual setting of variable t that indicates which model is actually analyzed.

- 5-7: The residual least squares fit is achieved by the Matlab-routine "fminsearch". It takes the optimizer setting defined in vector "opt" and the parameter start values defined in file "input.m" also as start values for the optimization. The routine calls the file "Model_Equation.m" in which for either model the model solution for the parameterization proposed by the optimization routine is calculated. This file of course also contains the definition of the optimization criterion that is calculated from the model solution and has to be optimized (in this case the residual least squares that have to be minimized).
- 9-15: The optimum parameterization that minimizes the residual least squares, as well as the corresponding residual least squares value, are saved to the workspace.

Model_Equation.m:

In this file for either model the model solution for the parameterization proposed by the optimization routine is calculated. This file of course also contains the definition of the optimization criterion that is calculated from the model solution and has to be optimized (in this case the residual least squares that have to be minimized). This file calls the file "Start_ODE_Solver_MM.m " (see above), "Start_ODE_Solver_Moser.m" (see above) respectively, by which the model solution for a model parameter vector V at the measuring points defined in file "input.m" is calculated for the model actually analyzed.

- 11: Calculation of the model solution for the Michealis-Menten model for the parameter vector V at the measuring points defined in file "input.m".
- 13-15: Definition and calculation of the residual least squares as optimization criterion for the fit of the Michealis-Menten model.
- 18: Calculation of the model solution for the Moser model for the parameter vector V at the measuring points defined in file "input.m".
- 20-22: Definition and calculation of the residual least squares as optimization criterion for the fit of the Moser model.

Start_ODE_Solver_MM_Plot.m:

This file is called to calculate the model solutions "y_int" for a custom defined interval "x_int" at a model parameterization "V". The difference to file "Start_ODE_Solver_MM.m" is, that the latter provides only the model solutions for the experimental measuring points defined in file "input.m" while the former provides the number of solutions that is required for a appropriate plot.

Start_ODE_Solver_Moser_Plot.m:

Analogous to file "Start_ODE_Solver_MM_Plot.m".

Posterior_Ordinate_Growth_Model.m:

In this file estimates of the posterior ordinate as well as the log prior probability and log likelihood of the candidate vector are calculated which are essential for the estimation of the log marginal likelihood. This is achieved for each model according to the method proposed by Chib and Jeliazkov. Log probabilities are applied as far as possible for computational reasons as explained for "Data_Analysis.m". Extremely low likelihood values are set to zero at a certain level by the program automatically and thus cannot be logarithmized anymore yielding a NaN ("not a number") entry that causes a breakdown of the computational process. Thus all zero likelihood values are set to 10^{-300} which is a value that one the one hand still can be logarithmized but on the other hand is so small that it is still negligible for the computational result compared to the values from the considerable parts of the distribution which are relevant for the computational result.

- 3-9: Defining and loading necessary experimental input data.
- 12-16: Calculation of the model solutions for the candidate parameter vector by the files "Start_ODE_Solver_MM.m" and "Start_ODE_Solver_Moser.m", respectively (see explanation above). These solutions are necessary for calculation of the likelihood.
- 20-23: Definition of the size of the posterior parameter distribution sample and the random sample from the Gaussian distribution characterized by the applied covariance matrix and candidate vector.
- 27-37: Calculation of the log likelihood of the candidate vector
- 39-57: Calculation of the prior probability of the candidate vector.
- 59-104: Calculation of the numerator of the posterior ordinate estimate for each model.
- 60-102: In this loop for each parameter vector of the representative sample of the posterior distribution the prior probability is calculated as well as the probability of the candidate vector of being proposed in a Markov-Chain process if this parameter vector is the last accepted chain state (covariance matrix applied to the proposal density equals covariance matrix estimate of posterior probability distribution).
- 68-73: Calculation of the log prior probability of a sample vector for the Michaelis-Menten model.
- 75-76: Calculation of the referring probability of proposal as exposed above.
- 80-85: Calculation of the log prior probability of a sample vector for the Moser model.
- 87-88: Calculation of the referring probability of proposal as exposed above.
- 91-95: Calculation of acceptance probability of the candidate vector if the sample vector was the last accepted chain state and the proposal density equal to the covariance matrix applied to posterior ordinate estimation.
- 97: Calculation of the probability of move from sample to candidate vector in a Bayesian Markov-Chain which is the product of proposal and acceptance probability.
- 104: Calculation of the numerator by averaging the probability of move over the whole posterior parameter distribution sample.
- 108-171: Drawing a random sample from a Gaussian probability distribution with the candidate vector as expected value and the covariance matrix applied here for posterior ordinate estimation as distribution variance. As proposed by Chib and Jeliazkov, the sample size should be chosen equal to the size of the posterior parameter distribution sample. For each sample vector, calculation of the probability of this vector to be accepted in a Bayesian Markov-Chain if the last accepted state was the candidate vector. The acceptance probability is calculated from the referring prior probabilities and likelihoods.
- 114-142: Calculation of the log prior probability of the sampled vector for the model that is actually analyzed.
- 144-164: Calculation of the model solution for the sampled parameter vector and the referring likelihood value.
- 166-170: Calculation of the non-log acceptance probability. Probability values larger than 1 are set to 1 as a probability of more than 100% is impossible.
- 173: Calculation of the denominator as acceptance probability average over the whole sample.
- 174-176: Saving all important intermediate results to the workspace.
- 178: Calculation of the posterior ordinate.

Plot_Prior_Post_Para_Distr.m:

This function requires a representative sample of the posterior parameter probability distribution of the model to be analyzed. This can be provided by running function Mltpl_Chn_Rndm_Wlk_Cov_Updt_Growth_Model.m before or loading its output into the workspace. Plot_Prior_Post_Para_Distr.m produces a representative sample from the prior parameter probability distribution and produces a comparative graphical exposition of the prior and posterior probability distribution for all parameters and models. The output of that function is saved to the workspace and is needed for further analysis in by the function Bayesian_Model_Validation_Growth_Model.m. The function Plot_Prior_Post_Para_Distr.m offers two different opportunities to generate a representative sample from the prior parameter distribution. The first one is to sample from a certain Gaussian-Distribution which has been accepted as prior knowledge (e.g. a Gaussian-Distribution with a certain expected value (e.g. parameterization from a least-squares fit of the model to the data) and a deviation of 50% of the expected value as uncertainty). The second one is the repeated generation and model fit of artificial "experimental" data from the real experimental date by Monte-Carlo-Simulation (exp. data that could have occurred with respect to the known error of experimentation). By that a sample of hypothetic parameterizations is produced whose probability distribution can be applied as prior knowledge.

- 4: Definition of the model and parameter sample to be analyzed (to be adapted).
- 5: Calls the required input (initial data base) from the workspace.
- 6: Definition of the representative sample size to be drawn and analyzed.
- 8-198: This loop achieves the whole process (compare above) one after another for the two different models.

- 11-12: Has to be activated if the first opportunity of introducing prior knowledge (compare above) was applied (definition of distribution variance and expected value). Otherwise inactivate.
- 15-19: Calls the required posterior samples from the workspace.
- 21-28: A representative posterior sample of the required sample size defined before is created from the output of the various Markov-Chains.
- 32-61: Generation of the prior sample.
- 39-57: Has to be activated if the second opportunity of introducing prior knowledge (compare above) was applied. Otherwise inactivate.
- 40-52: Generation of simulated "experimental" data sets to be fit by the models.
- 54-57: Model fit to simulated "experimental" data. The model parameter vector is stored to the prior sample vector.
- 59: Has to be activated if the first opportunity of introducing prior knowledge (compare above) was applied (random sampling from the distribution defined in 11-12). Otherwise inactivate.
- 63: Sample has to be saved to the workspace for further use in Bayesian_Model_Validation_Growth_Model.m.
- 65-77: Definition of required vectors and measures.
- 78-186: For the model actually analyzed the prior and posterior distributions are calculated from the samples and are plot together into a single figure for each parameter.
- 80-85: Definition of posterior distribution expected value and variance. Generation of posterior random sample (estimation of distribution by assumption of Gaussian nature). These function lines are activated if a suitable plot from the posterior Markov-Chain sample was not possible. The Markov-Chain sample arranged in line 21-28 is overwritten here. Otherwise inactivate.
- 86-89: Calculation of minimum, maximum, mean and definition of sample analysis range for prior and posterior sample.
- 91-100: Sorting of prior sample events into the intervals of the analysis range. Calculation of mean and probability for each interval.
- 102-105: Calculation of prior cumulative probability for localization of 5% and 95 % quantile.
- 107-117: Localization of prior quantiles for plot.
- 120-131: Plot of figures for the prior probability distribution.
- 133-185: Processing of posterior sample analogous to processing of prior sample.
- 187-197: Saving data to be further used to the workspace.

Bayesian_Model_Validation_Growth_Model.m :

In this file the prior and posterior probability of experimental observations at defined process conditions as well as the model validation as proposed by Geweke is computed corresponding to the theory exposed in chapter 2.3.3.2.4. . This function requires for each model to be analyzed a representative sample from the prior and posterior parameter distribution. The prior sample is provided by function Prior_Parameter_Distribution.m or by the function described here (in case of a defined standard deviation and expected value of the parameter distribution assumed as prior knowledge). A representative sample from the posterior parameter distribution (output of Mltpl_Chn_Rndm_Wlk_Cov_Updt_Growth_Model.m) has to be saved to the workspace before this function can be used.

- 3-6: Loading the data required for analysis from Input_Growth_Model.m. The vector "para" has to be adapted to the model to be analyzed (input.Parameter_Startvalues{1,…}).
- 7-9: In the vector C the experimental data is stored to ease the further programming.
- 11: Here the process variable settings are defined for which the probability of experimental observations has to be calculated from the Bayesian parameterized models. The whole range of settings has to be chosen here (X=[1:length(Cx)]) to achieve the model validation.
- 13: The size of the sample to be drawn from the prior and posterior distribution has to be defined. To produce a representative sample at least a size of 10^4 is proposed. This choice significantly influences the computing time.
- 15-387: This loop achieves all analytic tasks (probability distribution of observations, model validation) for the prior (t=1) and posterior (t=2) probability.
- 17-26: In case of posterior analysis the posterior parameter sample of the model to be analyzed is uploaded from the workspace and sorted into P_S to ease further programming and computing. Adapt line 18 according to the model to be analyzed! .

- 32-63: For each parameter vector of the parameter sample to be analyzed the model solutions of the model to be analyzed are calculated at the process settings defined in 11.
- 34-47: In case of analysis of the prior distribution a sample can be loaded from the workspace (e.g. output from Plot_Prior_Post_Para_Distr.m) or is produced here by random sampling from a prior distribution with "para" as expected value and defined standard deviation (both defined in line 41). Inactivate / activate the program lines according to the method chosen.
- 51-54: Calculation of model solutions. Activate the model to be analyzed!
- 56-58: Sorting out the model solutions at the process settings to be analyzed.
- 60-61: Saving the model solution separately for cell and substrate concentration for each parameter vector of the parameter sample.
- 70-385: In this loop the calculation of the probability distribution of observations at the process settings defined above as well as the model validation is achieved here either for the cell (t=1= or the substrate (t=2) concentration.
- 75-189: For the process variable settings to be analyzed the probability distribution of observations is calculated (the actual process variable, distribution and model have been defined in the higher-ranking loops). Inactivate this loop if only the model validation should be achieved.
- 77: For the process setting actually analyzed the solutions for each parameter vector of the parameter sample are sorted into vector "y_vec".
- 78-92: The spectrum of observations to be analyzed for the possibility of occurrence is defined. The minimum and maximum value has to be adjusted by hand for each figure / analysis to gain a significant plot / exposition of the distribution.
- 96-106: For each interval of the observation spectrum the probability that this value can be experimentally observed is calculated for each solution in "y_vec" with respected to the error of experiment repetition.
- 112-119: The relative probability of occurrence is calculated for each interval value.
- 121-146: The position of 5%- and 95%-Quantile as well as the position of experimental and most probable value in the observation spectrum analyzed are calculated.
- 148-187: Plot of the probability distribution of observations for the distribution, model and process variable actually analyzed (defined by the higher-ranking loops). Including plot of quantiles, exper. and most probable value.
- 191-383: Computation of model validation for the distribution, model and process variable actually analyzed (calculation for the process settings defined in line 11; choose all process settings exp. analyzed here for an optimal validation).
- 197-212: Vector "Y_Vec" contains for each sampled parameter vector (column) the model solutions (row) (with respect to the distribution, model and process variable actually analyzed (defined by higher-ranking loops)). For each of these model solutions a sample of theoretically occurring observations is generated by random sampling with respect to the error of experiment repetition / measurement. Mean and variance are calculated for each sampled observation.
- 199-211: Generation of a sample of size "Smpl_Sz" of theoretically occurring observations from a defined observation. Calculation of mean and variance for each sampled observation.
- 202-207: For each process setting of a defined model solution a theoretically occurring observation is calculated separately with respect to the error of exp, repetition / measurement.
- 213-225: Calculation of minimum and maximum of the solution means and variances calculated in line 197-212. Definition of the spectrum / interval of the mean and variance values for further analysis.
- 226-247: Sorting of all theoretically occurring mean and variance values (calculated / sampled in line 199-211) into the corresponding spectrum intervals. Calculation of relative probabilities for the intervals.
- 249-250: Calculation of mean and variance of the experimental observed data.
- 252-262: Calculation of the cumulative probability distribution for the further localization of quantiles.
- 264-279: For the observation mean: Localization of the position in the relative probability distribution for the 5%- and 95%-Quantiles as well as for the experimental values.
- 281-321: Plot of the relative probability distributions including quantiles and exp. values for model validation.
- 324-339: For the observation mean: Localization of the position in the relative probability distribution for the 5%- and 95%-Quantiles as well as for the experimental values.
- 341-381: Plot of the relative probability distributions including quantiles and exp. values for model validation.

Bayesian fit, discrimination and analysis of SPD models

In the following the m-files of the program are described in the order they are applied to the computational process. The numbers indicate the line of the specific m-file currently explained.

Mltpl_Chn_Rndm_Wlk_Cov_Updt_KWD.m:

With this file the whole computational process of the Bayesian method applied here is started. This file provides estimates of the expected value (vector "mean_vec") as well as the variance (vector "var_vec") of the model parameter probability distribution at defined steps during the convergence of this distribution towards the posterior. These estimates are needed to graphically expose the convergence process and to identify the point of chain convergence. This file thus also provides a representative sample of parameter vectors from the posterior distribution which is taken from the parallel Markov-Chains after convergence (vector "Par_Smpl"). Additionally a vector "lik" is provided that contains the likelihood values of the parameter vectors contained in the vector "Par_Smpl". The latter two vectors are applied to the calculation of expected value and variance of the posterior model parameter distribution as well as the relative model probability for model discrimination. The output of this file which is basically the output of the Bayesian Markov-Chain process is saved to the workspace. It is necessary for further data analysis by Bayesian means in the file "Data_Analysis.m" which loads the data from the workspace automatically. It is recommended to save the output of this file on a data storage medium. In this case a Bayesian model fit, analysis and discrimination by file "Data_Analysis.m" can be achieved without running the whole Markov-Chain process again which is extremely time consuming.

- 5-6: Calls file "Input_KWD.m" to introduce the basic data input necessary for the further computational process. In the following this data is defined as data matrix "input" which is assigned to the workspace and is globally valid / accessible to all associated m-files.
- 7-10: Here the process variables of the Bayesian Markov-Chain-Monte-Carlo process as well as its specific acceleration method are set up. These variables are the number of updates / re-estimates of the covariance matrix of the proposal density, the number of tunes of the covariance matrix of the proposal density within each re-estimation cycle of that matrix, the number of steps of the Markov-Chain within each tuning cycle of that matrix and finally the number of parallel Markov-Chains.
- 12: In the following the variable "t" defines the model treated. For model discrimination both models (t=1:2) have to be analyzed by Bayesian Markov-Chain-Monte-Carlo yielding the output mentioned above for each model.
- 14-18: Vector "Param" contains the start values for the parallel Markov-Chains. In the following the vector "Cov_Mat" takes up the re-estimated covariance matrix of the proposal density which is updated in each cycle of the loop initiated in line 20. The vectors "mean_vec" and "var_vec" take up the estimates of expected value und variance of the converging model parameter distribution as exposed above. The vector "par_tot" takes up the parameter vectors for each step of all parallel Markov-Chains. Except the last step of the preceding cycle this vector is cleared after each covariance re-estimation cycle to safe CPU and memory. In the last cycle (for which chain convergence has to be assured) the content of vector "par_tot" is saved as "Par_Smpl" (line 60) which is needed for further computation as exposed above.
- 20: This loop starts 1000 parallel Markov-Chains with 20 update cycles of the covariance matrix of the proposal density. Each update cycle consists of 10 tuning cycles of that matrix and each tuning cycle of 20 Markov-Chain steps. Thus each Markov-Chain consists of 10.000 steps in total. Covariance matrix update and tuning is implemented here.
- 25: Here m-file "Mltpl_Chn_Rndm_Wlk_KWD.m" is called which achieves 10 tuning steps of the covariance matrix of the proposal density within each update cycle. Within in each tuning cycle of file "Mltpl_Chn_Rndm_Wlk_KWD.m" 20 steps are achieved for each of the parallel Markov-Chains. File "Mltpl_Chn_Rndm_Wlk_KWD.m" provides the vector "par_tot" (see above) and the vector "lik" that contains the referring likelihood values for the parameter vectors in "par_tot".
- 27-54: m-file "Mltpl_Chn_Rndm_Wlk_KWD.m" provides the output of 10 tuning cycles each consisting of 20 steps for each of the 1000 parallel Markov-Chains. The last 20 steps for each of the 1000 parallel Markov-Chains are used to estimate the expected value and covariance which represent an estimation of the current state of the parameter probability distribution converging to the posterior distribution. After each re-estimation cycle these distribution parameters are saved in the vectors 'mean_vec' and 'var_vec' for further use (see above).
- 56-58: The actual variance of the parameter probability distribution is applied to the covariance matrix of the proposal density for the next re-estimation cycle.
- 60-65: The vector 'par_tot' that contains the parameter vectors of all parallel Markov-Chains of each entire re-estimation cycle is cleared (saving CPU) except the last vector of each chain which is required as starting point for the next re-estimation cycle. For the last re-estimation cycle (completed convergence has to be assured by graphical analysis) the parameter vectors for all parallel Markov-Chains are saved in the vector "Par_Smpl" which is at completed convergence a representative sample of the required posterior parameter probability distribution.

- 69-105: Plot of the convergence of the parameter probability distribution towards the posterior distribution. Plots are achieved for both models and each of the characteristic distribution parameters (expected value and variance).
- 109-119: Saving all output of the Markov-Chain process that is required for further analysis to the workspace.

Mltpl_Chn_Rndm_Wlk_KWD.m:

- For each re-estimation cycle m-file "Mltpl_Chn_Rndm_Wlk_KWD.m" achieves 10 tuning steps of the covariance matrix of the proposal density with each step consisting of 20 steps each for 1000 parallel Markov-Chains. The procedure of running parallel Markov-Chains with iterative tuning of the proposal density is implemented here. File "Mltpl_Chn_Rndm_Wlk_KWD.m" provides the vectors "par_tot" and "lik" whose further function has been already exposed above.
- 5: This loop achieves 10 tuning steps of the covariance matrix of the proposal density each being based on 20 steps of 1000 parallel Markov-Chains each.
- 10: File "Rndm_Wlk_KWD.m" achieves 20 steps of 10000 parallel Markov-Chains each. This file provides for each tuning cycle the vectors "par", "Acc_Prob_Vec" and "Lik" which contain for each Markov-Chain the parameter vectors and the referring acceptance probability and likelihood value.
- 12-23: After each tuning cycle achieved by file "Rndm_Wlk_KWD.m" the latest chain steps achieved in that file are added to vector "par_tot" for each of the parallel Markov-Chains.
- 25: In the context of line 12-23 the referring likelihood values of the parameter vectors, chain steps respectively, are added analogously.
- 27-39: Calculation of the average acceptance probability of the additional chain steps generated in the latest tuning cycle by file "Rndm_Wlk_Growth_Model.m". Tuning of the covariance matrix of the proposal density that applies the average acceptance probability as decision criterion.

Rndm_Wlk_KWD.m:

- For each tuning cycle of file "Mltpl_Chn_Rndm_Wlk_KWD.m" file "Rndm_Wlk_KWD.m" achieves 20 steps for 1000 parallel Markov-Chains. As output this file provides the vectors "par", "Acc_Prob_Vec" and "Lik" whose function and further processing already has been exposed above.
- 3-8: The input data required for the Bayesian Markov-Chain process is uploaded.
- 10-12: The vectors taking up the process output subsequently are defined.
- 14-149: This loop achieves 1000 Markov-Chains one after another. Within that loop 20 chain steps are achieved.
- 16-27: Definition of the start vector for each chain. If the Markov-Chain has to be started (first re-estimation cycle and first tuning cycle) the input start vector is applied. Otherwise the last vector of the chain to be elongated (last vector of the last cycle) is applied as start vector. As all chains have to be subject to identical conditions for the reasons all chains are started (first re-estimation cycle and first tuning cycle) from the same start vector defined in the file "input.m".
- 31-145: This inner loop achieves 20 chain steps for each chain called up by the outer loop (line 14-149, see above).
- 33-76: In this statement for each chain step the latest accepted state / parameter vector of the chain actually treated is defined, as well as the proposed new state / parameter vector of that chain.
- 35-55: For the first step of a chain (N=1) the latest accepted state is the start vector defined in line 16-22. Regarding the proposal of a new parameter vector as potential new chain state two situations have to be distinguished for the first chain step (N=1): The first situation (line 35-44) is that the program is still in the first covariance estimation cycle (a=1). Here no covariance of the actual state of the converging parameter probability distribution has been estimated so far as this is done at the end of each estimation cycle (see Mltpl_Chn_Rndm_Wlk_Cov_Updt_KWD.m line 27-58). In this case the proposed candidate vector is initially sampled randomly from a Gaussian distribution with the start vectors defined in file "input.m" as mean and a standard deviation of 50% (line 39-42). This procedure is chosen to initially provide a broad parameter range for the Bayesian Markov-Chain process within which the random process can narrow the investigated parameter range to a coarse estimate of the range of values that are most possible in the light of the experimental data. The second situation (line 44-53) is that the program is in the second or an upper re-estimation cycle and thus a covariance matrix of the actual state of the converging parameter probability distribution has been already estimated (see Mltpl_Chn_Rndm_Wlk_Cov_Updt_KWD.m line 27-58). The proposed parameter vector is in this case generated by random sampling from a Gaussian distribution with the latest estimated covariance matrix as variance and the start vector (compare line 16-22) as mean.
- 55-76: For the second and all following chain steps (N>1) two situations can be distinguished: The first situation (line 57-66) is that the program is still in the first covariance estimation cycle (a=1). For this case the same procedure for proposal of a candidate parameter vector is valid as exposed for the first chain step (N=1) above. The second situation (line 66-74) is that the program is in the second or an upper covariance re-estimation cycle. In this case the latest accepted state / parameter vector from a

preceding chain step is defined as origin for the proposal of a candidate parameter vector to be accepted or rejected as next chain state. The proposed candidate is derived by random sampling from a Gaussian distribution with the latest accepted chain state as mean and the latest estimate of the covariance matrix as variance.

- 78-117: Calculation of prior probability and likelihood of the actual state and the parameter vector that has been proposed as next candidate. As prior knowledge about model parameterization a Gaussian distribution with the start vectors defined in file "input.m" as mean and a standard deviation of 50% is applied. The choice of such a rather small prior knowledge (instead the results of a residual least squares fit could be applied as improved prior knowledge about the mean combined with a standard deviation that has been calculated from a repeated residual least squares fit of the model to data that has been generated by repeated random addition of noise to the experimental data) has been made to maximize the impact of the experimental data on the posterior distribution. All probabilities are calculated here as log values for the reasons exposed in the beginning of the description of file "Posterior_Ordinate_KWD.m" below.
- 119-143: Implementation of the criterion and decision on candidate vector acceptance or rejection. The formula for the calculation of the criterion is converted here to calculate a log value for the criterion from the log probabilities. That log value is re-converted into a non-log value that is then applied for the acceptance-rejection-decision.
- 147: Saving the chain of accepted chain states / parameter vectors for the chain that has been actually treated by the loop from line 14-160.

Input_KWD.m :

This file contains all input data on the models to be analyzed as well as the experimental data required for the computational process.

- 3-10: Experimental data for the target values acid value, rancimat and tocopherol at the different process variable settings of the SDoEs applied here
- 12-22: Definition of the process variable settings of the SDoEs applied here.
- 25-31: Values of the model parameterizations which are used as start values for the Markov-Chain process applied to the Bayesian determination of the most probable parameterization.
- 33-44: Expected value of prior parameter probability distribution.
- 42-44: Error / standard deviation of experimental analysis.

Start_Model_KWD.m:

- 5-7: Definition and uploading of input data required for calculation of model solutions.
- 12-28: Definition of 1st and 2nd grade models for the target values. Calculation of model solutions for all process variable settings analyzed.

Data_Analysis.m:

This file achieves the Bayesian model analysis and discrimination based on the output of file "Mltpl_Chn_Rndm_Wlk_Cov_Updt_KWD.m" which is the representative parameter sample from the posterior distribution and associated likelihood values as well as mean and variance values from the complete convergence process of the Markov-Chain.

- 5-182: Estimation of the log marginal likelihoods for both models. The computations described in the following are thus just achieved for the model actually analyzed which is indicated by the value of variable t.
- 7-22: The input data is defined and loaded upon which the Bayesian analysis achieved here is based (see "input.m" above).
- 24-29: From the variances, which have been calculated as estimate for the posterior from the converged Markov-Chain, an average is calculated as final estimate for the posterior probability distribution. The parameter G defines the number of values from the chain end to be taken for averaging. It is important to only apply values from the converged chain. For that purpose the point of convergence has to be identified before graphically.
- 30-38: It is analyzed in this work if differences in the number of parameters between two models cause dimensionality problems in model discrimination. For this purpose the 1st grade model is converted into the same structure as the second grade model by adding additional model terms which are characterized by parameter distributions with expected values and variances of value zero. For computational reasons the values have to be set to negligible values close to zero. The vectors which store the variances and expected values of model parameterization are adapted according to that proposed procedure. Activate the corresponding lines.

- 40-78: For each parameter vector of the posterior sample the numerator of the posterior probability estimate (the denominator is a normalizing constant that is equal for each parameter vector) is calculated to identify the parameterization with the highest posterior probability (mode). This procedure accounts for the fact that in case of distribution non-symmetry the distribution expected value is not identical with the mode. The latter is applied as Bayesian model parameterization and candidate vector for the posterior ordinate estimation. In case of analysis of dimensionality problems in model discrimination (as exposed above for line 30-38) the parameter vector of the 1st grade model has to be adapted to the 2nd grade structure (activate corresponding lines).
- 80-86: Saving mode and variance estimate of the posterior distribution to the workspace.
- 88-155: Plot of the convergence of the parameter probability distribution towards the posterior distribution. The convergence plots are achieved for mean and variance as basic distribution parameters and for both models.
- 157-165: For both models the process variable setting is calculated at which the target value is maximized or minimized.
- 167-168: Definition of covariance matrix and candidate vector for posterior ordinate estimation.
- 170: Call of file "Posterior_Ordinate_KWD.m". This file achieves the calculation of the posterior ordinate as well as the log prior probability and log likelihood of the candidate vector. To ensure a high precision of the marginal likelihood estimate a parameter sample from the posterior distribution has to be applied that is sufficiently representative. Thus the whole representative sample taken from the Bayesian Markov-Chain output is applied for the posterior ordinate estimation. As far as possible by means of mathematical calculation rules all computational procedures in the course of marginal likelihood estimation are achieved by use of log probabilities. This approach is necessary as especially in case of likelihood calculations extremely low probabilities can occur which are not computational manageable anymore by the program.
- 174: Calculation of the log marginal likelihood from the output of "Posterior_Ordinate_Growth_Model.m".
- 176-180: Saving the log marginal likelihoods of the models to the workspace.
- 182: End of log marginal likelihood estimation loop.
- 184-185: Loading of the log marginal likelihood estimates for final calculation of the relative model probability.
- 186-188: Calculation of the relative model probability and saving to the workspace.

Func_KWD_x_opt.m:

Contains the model functions which are applied by the Matlab-routine "fminsearch" for the calculation of the function minimum.

- 3: Definition of the model parameterization. The mode estimate of the posterior parameter distribution (Bayesian model parameterization) is applied.
- 5-11: The Matlab-routine "fminsearch" also proposes process variable settings which are outside the range of normalized process variable settings investigated. In this case the proposals are limited to the minimum / maximum of the range.
- 13-24: Process models of 1st and 2nd grade for the target values of interest. In case of analysis of the above mentioned dimensionality problems in model discrimination, a model of 2nd grade structure has to be also activated for the 1st grade model (compare models for rancimat).
- 26-32: Saving proposed process variable setting and corresponding function value to the workspace. In each optimization cycle of the Matlab-routine "fminsearch" the value saved so far in the previous cycle is overwritten by the actual one. Thus at the end of the optimization procedure just the optimum process variable setting and corresponding function value remains in the workspace. This procedure is chosen as the last proposal and output of the function "fminsearch" in line 165 of Data_Analysis_KWD.m is not limited to the normalized process variable range analyzed and thus might exceed that range.

Posterior_Ordinate_KWD.m:

In this file estimates of the posterior ordinate as well as the log prior probability and log likelihood of the candidate vector are calculated which are essential for the estimation of the log marginal likelihood. This is achieved for each model according to the method proposed by Chib and Jeliazkov. Log probabilities are applied as far as possible for computational reasons as explained for "Data_Analysis.m". Extremely low likelihood values are set to zero at a certain level by the program automatically and thus cannot be logarithmized anymore yielding a NaN ("not a number") entry that causes a breakdown of the computational process. Thus all zero likelihood values are set to 10^{-300} which is a value that one the one hand still can be logarithmized but on the other hand is so small that it is still negligible for the computational result (compared to the values from the considerable parts of the distribution which are relevant for the computational result). (Example: Line 27-29.)

- 4-9: Defining and Loading necessary experimental input data.

- 11: Calculation of the model solutions for the candidate parameter vector by the file "Mod_Sol_KWD.m".
- 15-18: Definition of the size of the posterior parameter distribution sample and the random sample from the Gaussian distribution which is characterized by the applied covariance matrix and candidate vector. The choice of the sampled distribution and the sample size is implemented.
- 20-32: Calculation of the log likelihood of the candidate vector
- 34-41: Calculation of the prior probability of the candidate parameter vector.
- 43-81: Calculation of the numerator of the posterior ordinate estimate for each model.
- 44-79: In this loop for each parameter vector of the representative sample of the posterior distribution the prior probability is calculated as well as the probability of the candidate vector of being proposed and accepted in a Markov-Chain process if the parameter vector from the posterior sample is the last accepted chain state (the proposal density has to be equal to the covariance matrix applied here to the marginal likelihood estimation (=variance of the posterior probability distribution)).
- 47-48: Definition of the posterior sample vector to be analyzed in the current cycle of the loop.
- 49-57: Transformation of the 1st grade model parameter vector to the 2nd grade model structure (in case of analysis of dimensionality problems in model discrimination).
- 58-63: Calculation of the log prior probability of the posterior sample vector.
- 65-67: Calculation of the referring probability of proposal as exposed above.
- 68-72: Calculation of the acceptance probability of the candidate vector if the sample vector was the last accepted chain state and the proposal density equal to the covariance matrix applied to posterior ordinate estimation.
- 74: Calculation of the probability of move from sample to candidate vector in a Bayesian Markov-Chain which is the product of proposal and acceptance probability.
- 81: Calculation of the numerator by averaging the probability of move over the whole posterior parameter distribution sample.
- 89-125: Drawing a random sample from a Gaussian probability distribution with the candidate vector as expected value and the covariance matrix applied here for posterior ordinate estimation as distribution variance. As proposed by Chib and Jeliazkov the sample size should be chosen equal to the size of the posterior parameter distribution sample. For each sample vector calculation of the probability of this vector to be accepted in a Bayesian Markov-Chain if the last accepted state was the candidate vector. The acceptance probability is calculated from the referring prior probabilities and likelihoods.
- 91-92: Random sampling of a vector.
- 93-96: Transformation of the 1st grade model parameter vector to the 2nd grade model structure (in case of analysis of dimensionality problems in model discrimination).
- 97-100: Calculation of the log prior probability of the sampled vector.
- 105-118: Calculation of the model solution for the sampled parameter vector and the referring likelihood value.
- 120-124: Calculation of the non-log acceptance probability. Probability values larger than 1 are set to 1 as a probability of more than 100% is impossible.
- 128: Calculation of the denominator as acceptance probability average over the whole sample.
- 129-131: Saving all important intermediate results to the workspace.
- 133: Calculation of the posterior ordinate.

Mod_Sol_KWD.m:

Calculates for each model parameter vector the model solution for a defined target value.

- 3: Definition of the model parameterization.
- 6-21: For a defined parameter vector and model the function value for each process variable setting is calculated.
- 7-20: Process models of 1st and 2nd grade for the target values of interest. In case of analysis of the above mentioned dimensionality problems in model discrimination, a model of 2nd grade structure has to be also activated for the 1st grade model (compare models for rancimat).

Plot_Prior_Posterior_Parameter_Distribution.m:

This function requires a representative sample of the posterior parameter probability distribution of the model to be analyzed. This can be provided by running function Mltpl_Chn_Rndm_Wlk_Cov_Updt_KWD.m before or loading its output into the workspace. Plot_Prior_Posterior_Parameter_Distribution.m produces a representative sample from the prior parameter probability distribution and produces a comparative graphical exposition of the prior and posterior probability distribution for all parameters and models. The output of that function is saved to the workspace and is needed for further analysis by the function Bayesian_Model_Validation.m. The function Plot_Prior_Posterior_Parameter_Distribution.m generates a representative sample from the prior parameter distribution by sampling from a certain Gaussian-Distribution which has been accepted as prior knowledge (e.g. a

Gaussian-Distribution with a certain expected value (e.g. parameterization from a least-squares fit of the model to the data) and a deviation of 50% of the expected value as uncertainty).

- 5: Calls the required input (initial data base) from the workspace.
- 6: Definition of the representative sample size to be drawn and analyzed.
- 8-163: This loop achieves the whole process (compare above) one after another for the two different models.
- 11-17: Drawing a random sample from the prior distribution with vector "para" as expected value and a standard deviation which has to be defined in line 13 in percent of the expected value.
- 20-24: Calls the required posterior samples from the workspace.
- 25-31: A representative posterior sample of the required sample size defined before is created from the output of the various Markov-Chains.
- 33-44: Definition of required measures and definition of vectors that take up the required results to be calculated from the distribution samples.
- 46-149: For the model actually analyzed the prior and posterior distributions are calculated from the samples and are plot together into a single figure for each parameter.
- 48-53: Calculation of mean, variance, maximum and minimum of the prior distribution. Definition of sample analysis range for the prior sample.
- 55-65: Sorting of prior sample events into the intervals of the analysis range. Calculation of mean and probability for each interval.
- 67-70: Calculation of prior cumulative probability for localization of 5% and 95 % quantile.
- 72-82: Localization of prior quantiles for plot.
- 84-95: Plot of figures for the prior probability distribution.
- 97-148: Processing of posterior sample analogous to processing of prior sample.
- 151-161: Saving data to be further used to the workspace.

Bayesian_Model_Validation.m :

In this file the prior and posterior probability of experimental observations at defined process conditions as well as the model validation as proposed by Geweke is computed. This function requires for each model to be analyzed a representative sample from the prior and posterior parameter distribution. The prior sample is provided by function Plot_Prior_Posterior_Parameter_Distribution.m or by the function described here (in case of a defined standard deviation and expected value of the parameter distribution assumed as prior knowledge). A representative sample from the posterior parameter distribution (output of Mltpl_Chn_Rndm_Wlk_Cov_Updt_KWD.m) has to be saved to the workspace before this function can be used.

- 3-5: Loading the data required for analysis from Input_KWD.m. The vectors "para" and "SZ_2" have to be adapted to the model to be analyzed.
- 7-22: Definition of the process variable settings which are subject to further analysis as well as the experimental observations at these process settings. Only activate these elements here which are actually subject to analysis!
- 24: Definition of the sample size of the prior and posterior sample to be applied to further analysis.
- 26-394: This loop achieves all analytic tasks (probability distribution of observations, model validation) for the prior (t=1) and posterior (t=2) probability.
- 28-37: In case of posterior analysis the posterior parameter sample of the model to be analyzed is uploaded from the workspace and sorted into P_S to ease further programming and computing. Adapt line 29 according to the model to be analyzed!
- 41-82: For each parameter vector of the parameter sample to be analyzed the model solutions of the model to be analyzed are calculated at the process settings defined in line 7-22.
- 43-56: In case of analysis of the prior distribution a sample can be loaded from the workspace (e.g. output from Plot_Prior_Posterior_Parameter_Distribution.m) or is produced here by random sampling from a prior distribution with "para" as expected value and defined standard deviation (line 50). Inactivate / activate the program lines according to the method chosen.
- 63-78: Calculation of model solutions. Activate the model to be analyzed!
- 80: Saving the model solution separately for each parameter vector of the parameter sample.
- 87-214: In this loop the calculation of the probability distribution of observations at the process settings defined above is achieved either for the cell (t=1) or the substrate (t=2) concentration.
- 89: For the process setting actually analyzed the solutions for each parameter vector of the parameter sample are sorted into vector "y_vec".
- 89-92: The spectrum of observations to be analyzed for the possibility of occurrence is defined. The minimum and maximum value has to be adjusted by hand for each figure / analysis to gain a significant plot / exposition of the distribution.
- 96-105: For each interval of the observation spectrum the probability that this value can be experimentally observed is calculated for each solution in "y_vec" with respected to the error of experiment repetition.

- 111-118: The relative probability of occurrence is calculated for each interval value.
- 120-152: The position of 5%- and 95%-quantile as well as the position of experimental and most probable value in the observation spectrum analyzed are calculated.
- 155-212: Plot of the probability distribution of observations for the distribution, model and process variable actually analyzed (defined by the higher-ranking loops). Including plot of quantiles, exp. and most probable value.
- 216-392: Computation of model validation for the distribution, model and process variable actually analyzed (calculation for the process settings defined in line 7-16; choose all process settings exp. analyzed here for an optimal validation).
- 218-237: Vector "Y_Vec" contains for each sampled parameter vector (column) the model solutions (row) (with respect to the distribution, model and process variable actually analyzed (defined by higher-ranking loops)). For each of these model solutions a sample of theoretically occurring observations is generated by random sampling with respect to the error of experiment repetition / measurement. Mean and variance are calculated for each sampled observation.
- 224-236: Generation of a sample of size "Smpl_Sz" of theoretically occurring observations from a defined observation. Calculation of mean and variance for each sampled observation.
- 239-244: Calculation of minimum and maximum of the solution means and variances calculated in line 222-237. Definition of the spectrum / interval of the mean and variance values for further analysis.
- 246-267: Sorting of all theoretically occurring mean and variance values (calculated / sampled in line 222-237) into the corresponding spectrum intervals. Calculation of relative probabilities for the intervals.
- 269-270: Calculation of mean and variance of the experimental observed data.
- 272-282: Calculation of the cumulative probability distribution for the further localization of quantiles.
- 284-299: For the observation mean: Localization of the position in the relative probability distribution for the 5%- and 95%-quantiles as well as for the experimental values.
- 301-336: Plot of the relative probability distributions including quantiles and exp. values for model validation.
- 339-354: For the observation variance: Localization of the position in the relative probability distribution for the 5%- and 95%-Quantiles as well as for the experimental values.
- 356-391: Plot of the relative probability distributions including quantiles and exp. values for model validation.

Annex C – Figures for 5.2.1.1

- Run 1: Application of the parameterization with the highest posterior probability as candidate vector for model discrimination
 - o Concergence of Markov-Chains (exemplary)
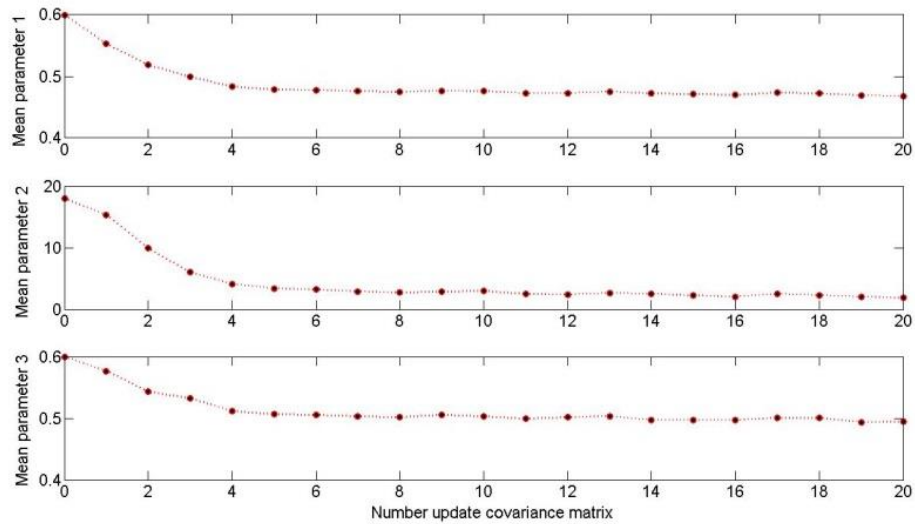
Figure C.9: Convergence mean MM-Model
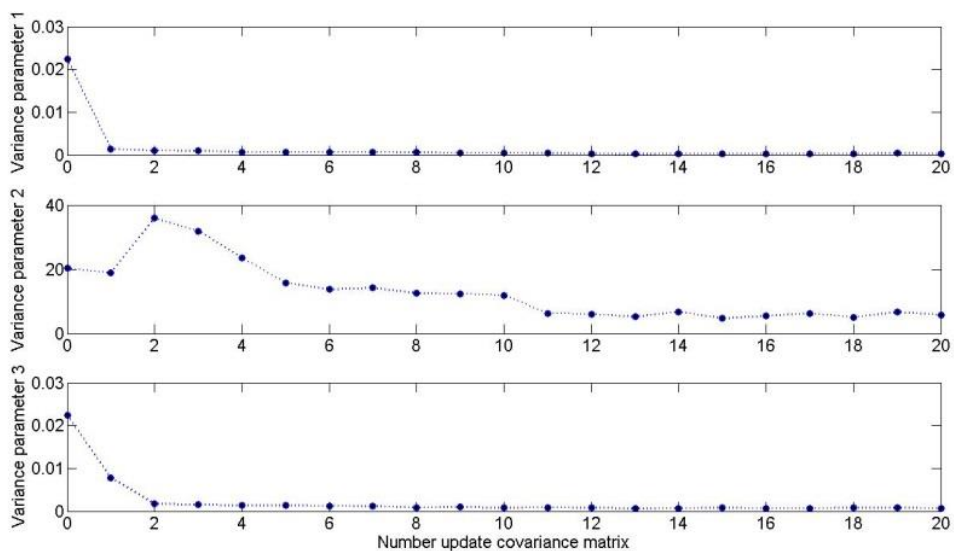
Figure C.10: Convergence variance MM-Model

144
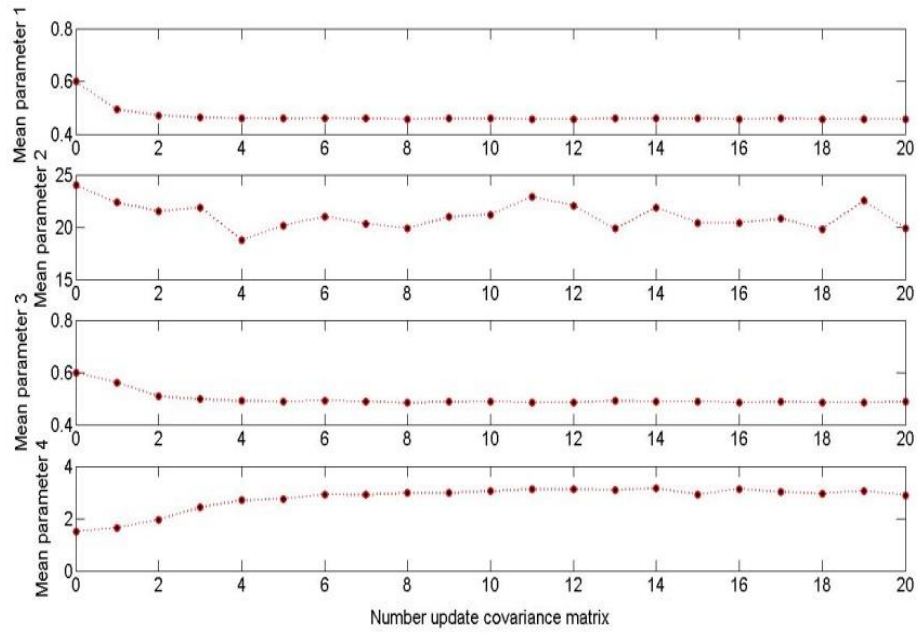
Figure C.11: Convergence mean Moser-Model
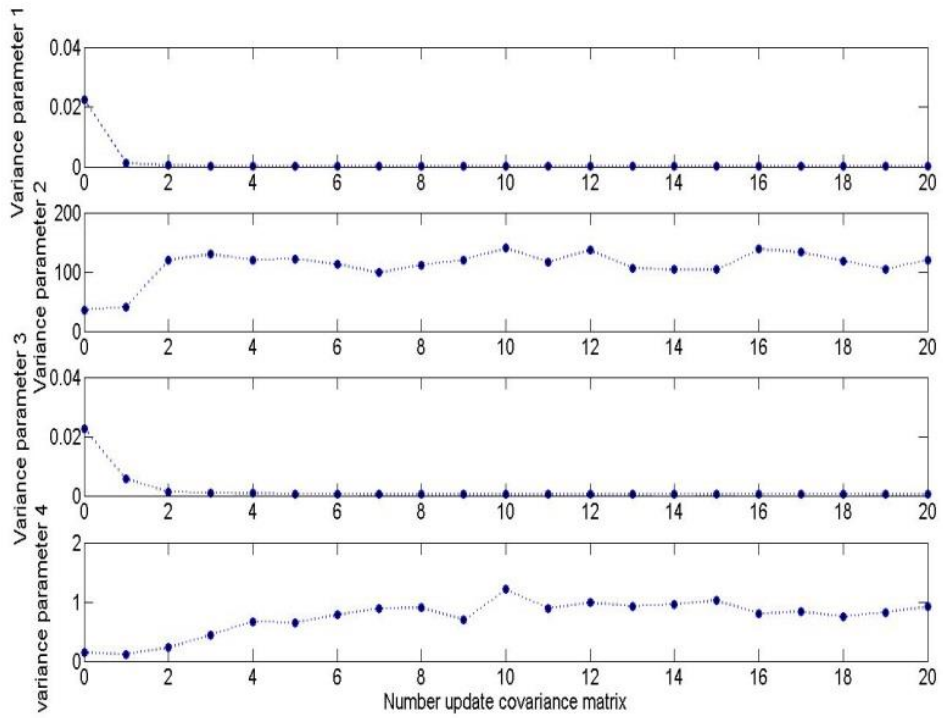
Figure C.12: Convergence variance Moser-Model

145
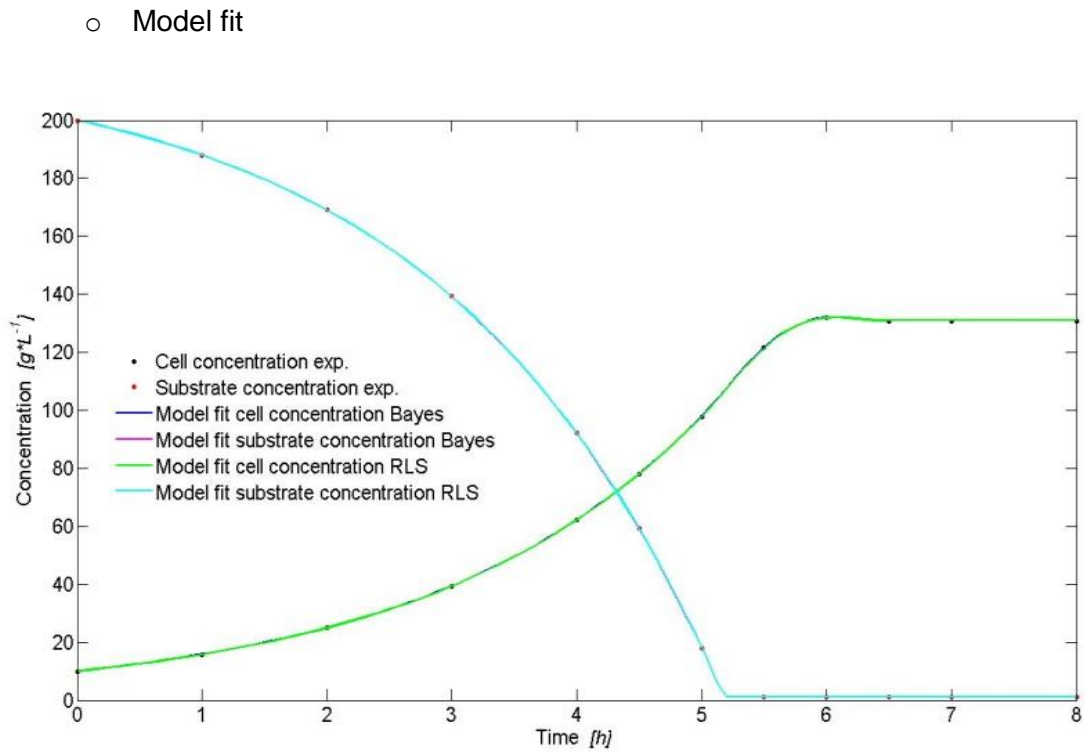
o Model fit

Figure C.13: Bayesian- and RLS-Model-Fit of MM-Model
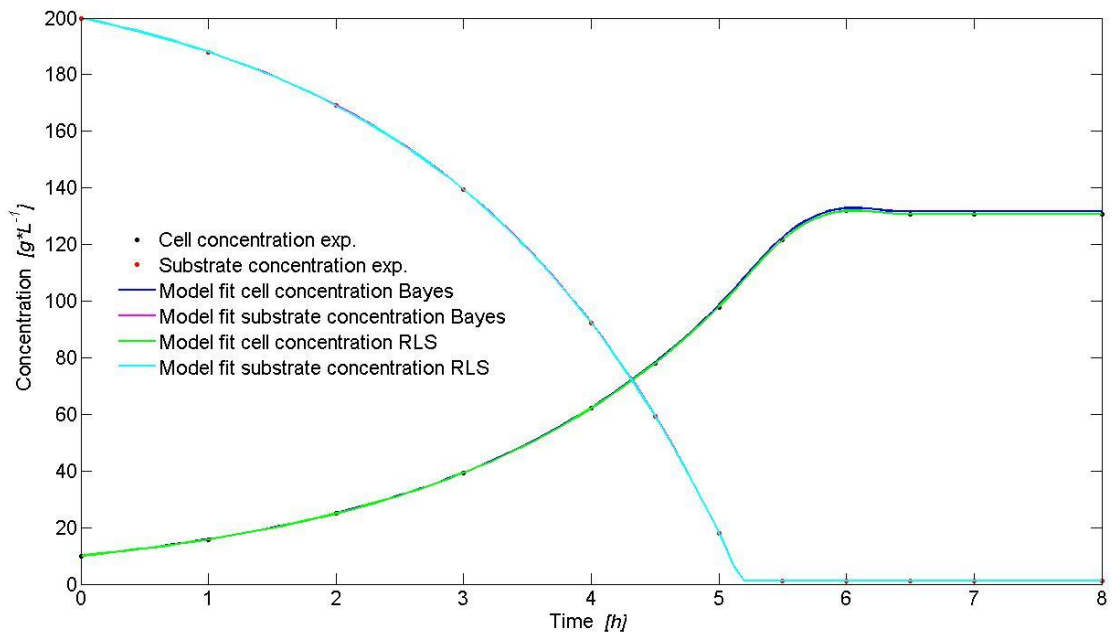
Figure C.14: Bayesian- and RLS-Model-Fit of Moser-Model

o Prior and posterior parameter distributions (exemplary)
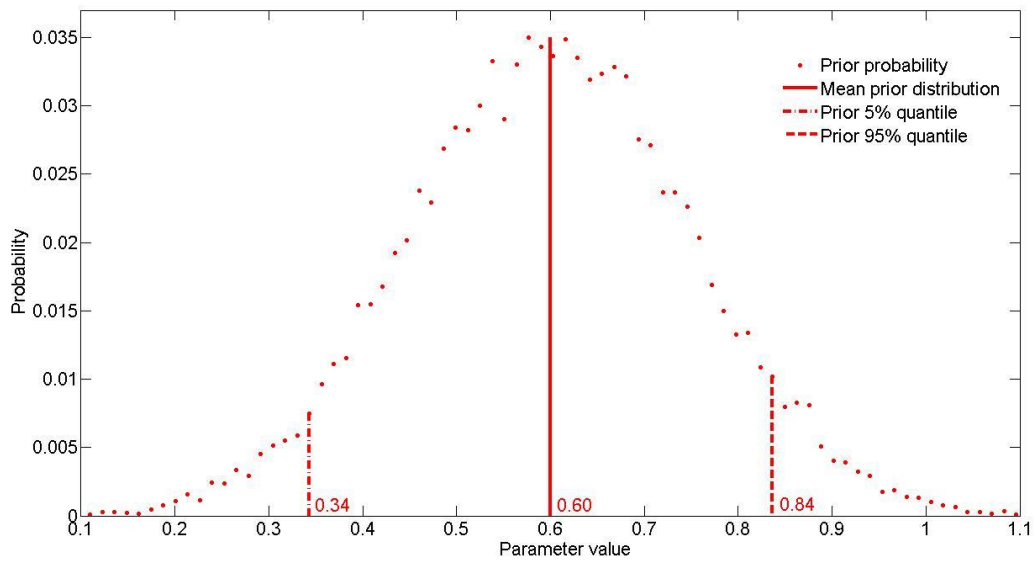
Figure C.15: Prior probability distribution of parameter 2 MM-Model
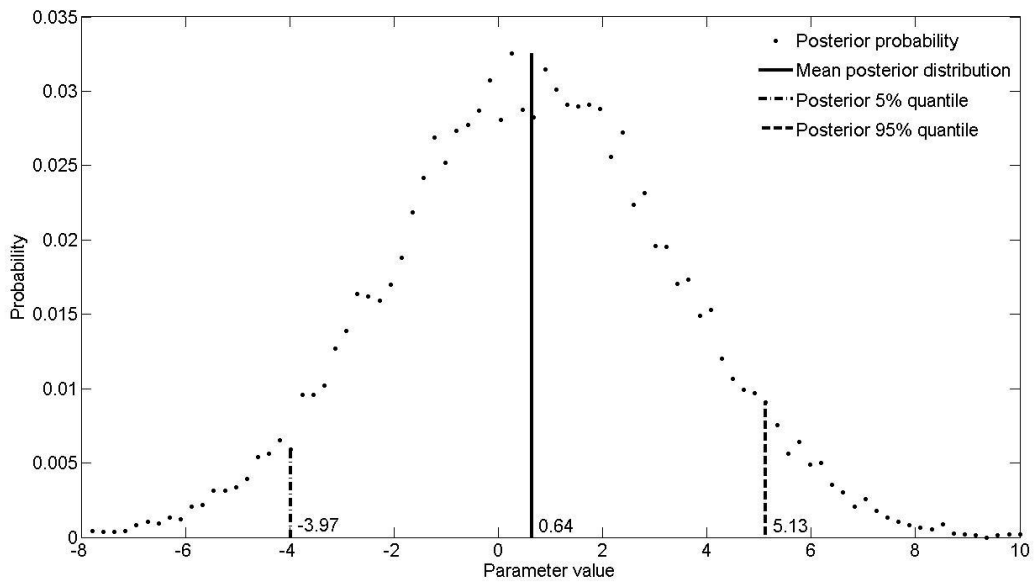
Figure C.16: Posterior probability distribution of parameter 2 MM-Model

147
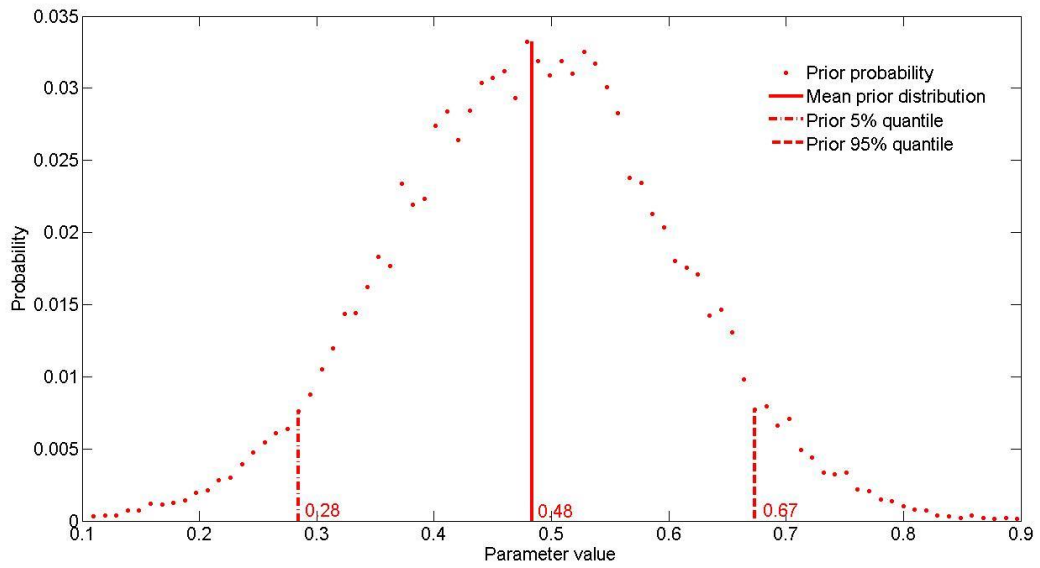
Figure C.17: Prior probability distribution of parameter 3 Moser-Model
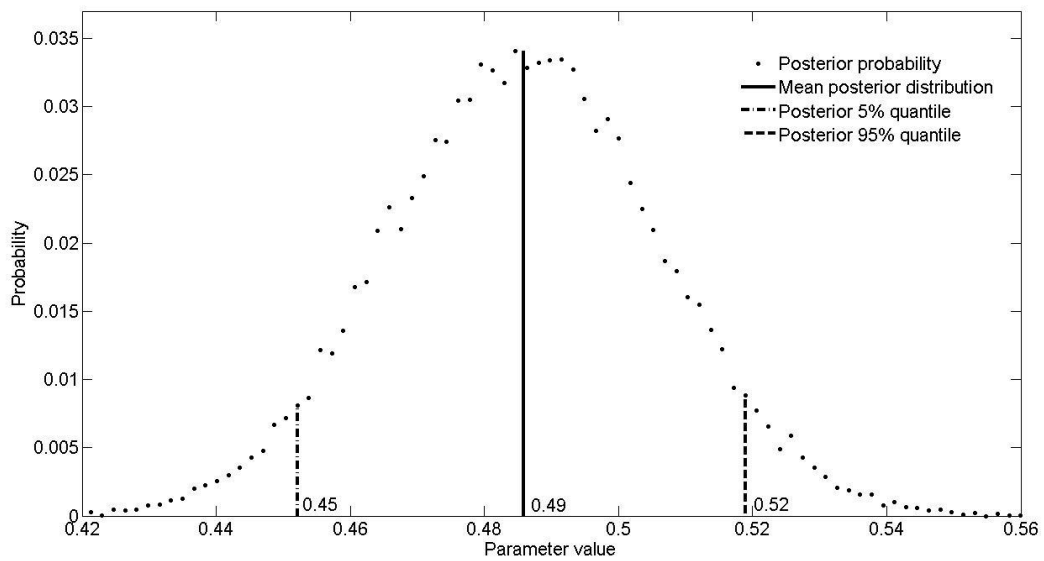
Figure C.18: Posterior probability distribution of parameter 3 Moser-Model
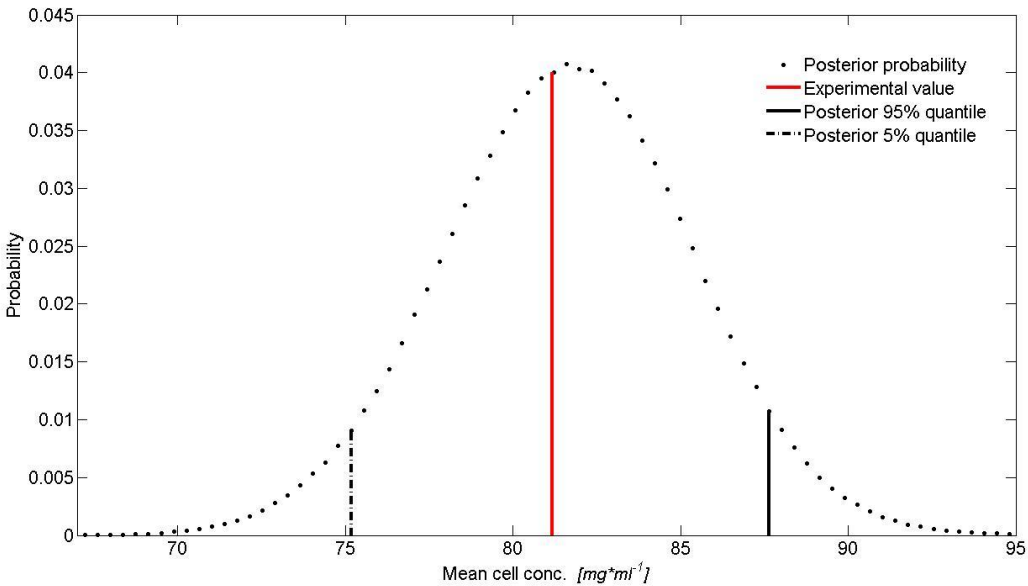
o   Model validation

Figure C.19: Validation mean cell concentration MM-Model
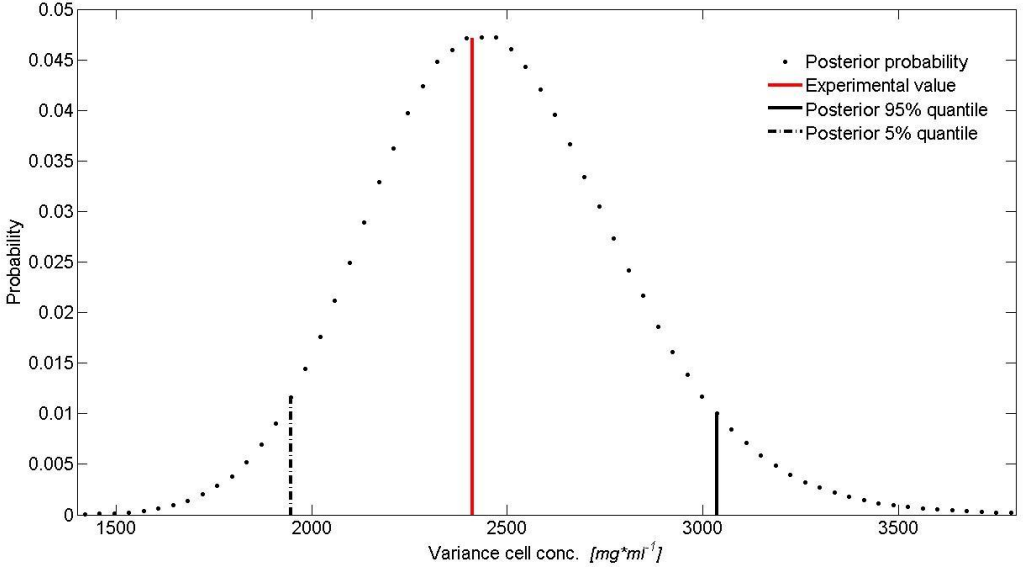
Figure C.20: Validation variance cell concentration MM-Model
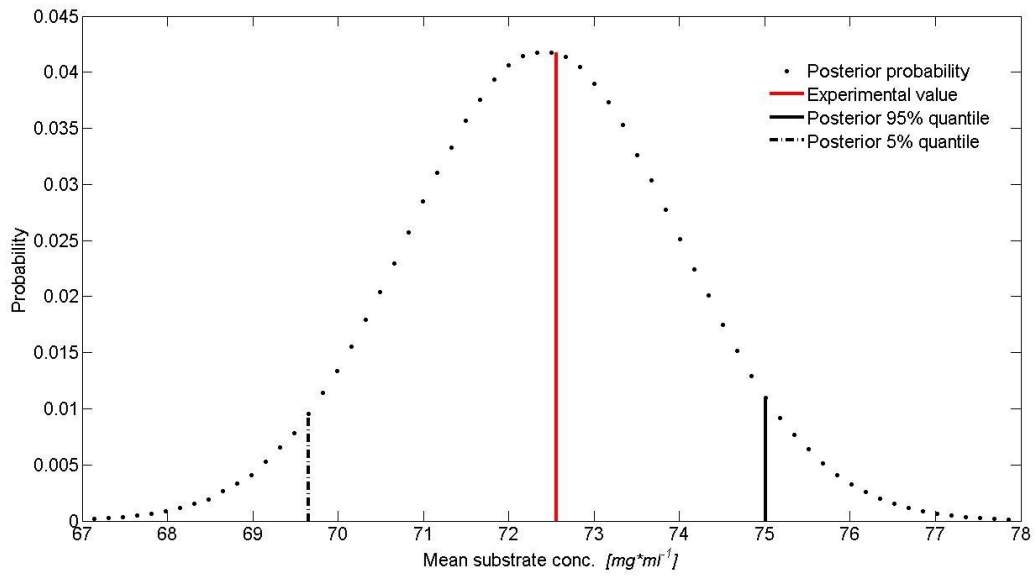
Figure C.21: Validation mean substrate concentration MM-Model
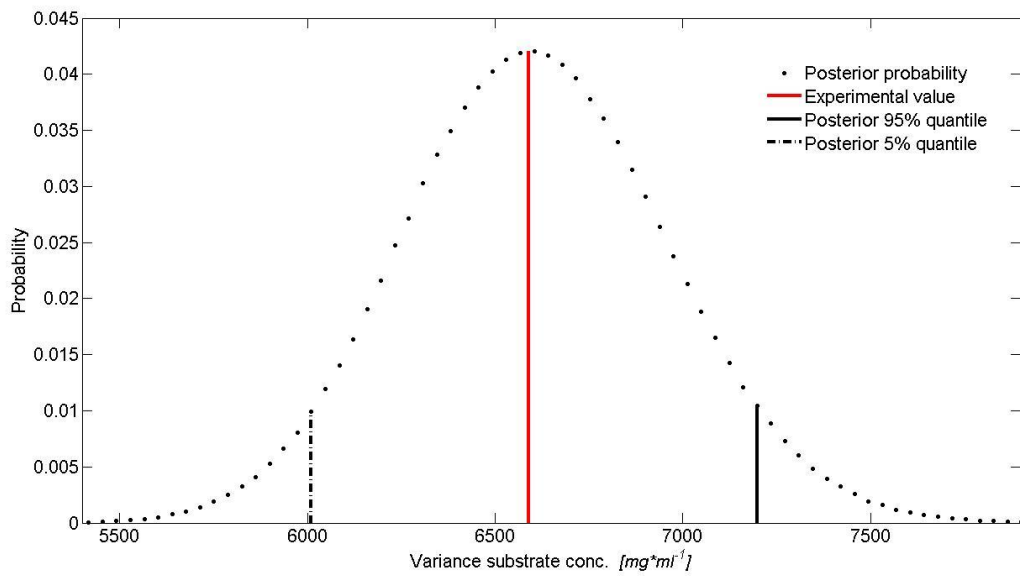
Figure C.22: Validation variance substrate concentration MM-Model
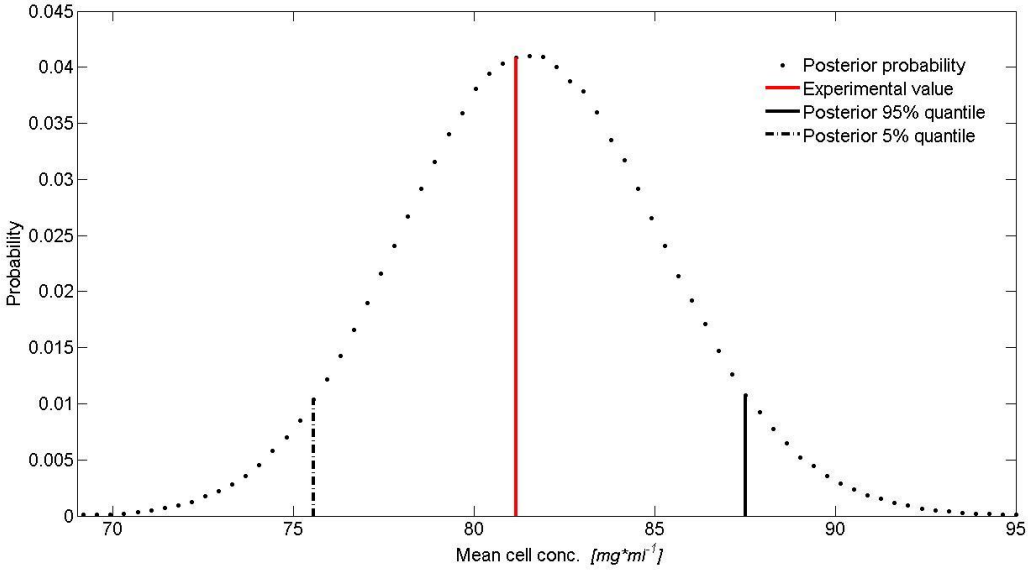
Figure C.23: Validation mean cell concentration Moser-Model
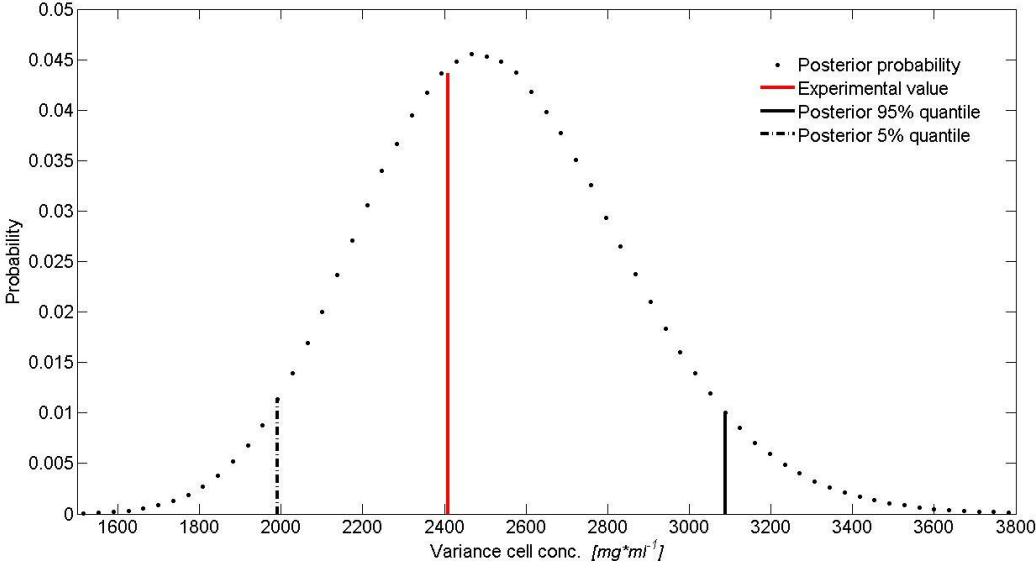
Figure C.24: Validation variance cell concentration Moser-Model
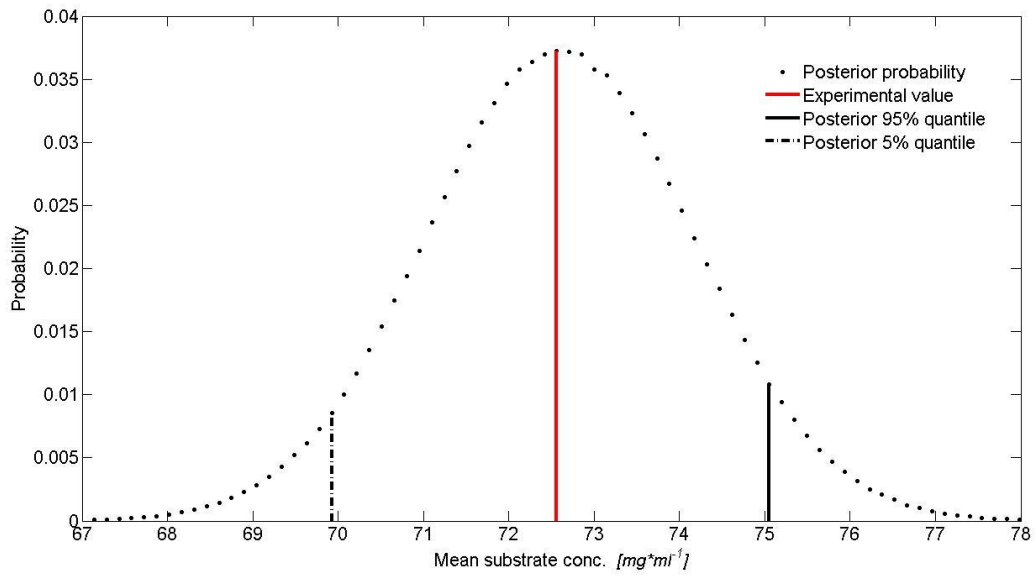
Figure C.25: Validation mean substrate concentration Moser-Model
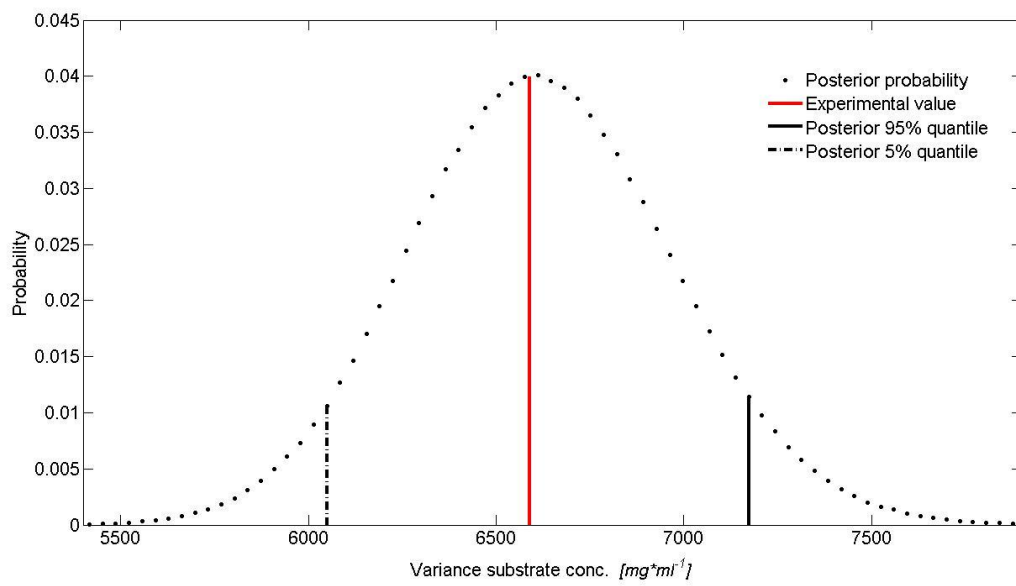
Figure C.26: Validation variance substrate concentration Moser-Model

152

o Posterior probability distribution of experimental observations (exemplary)
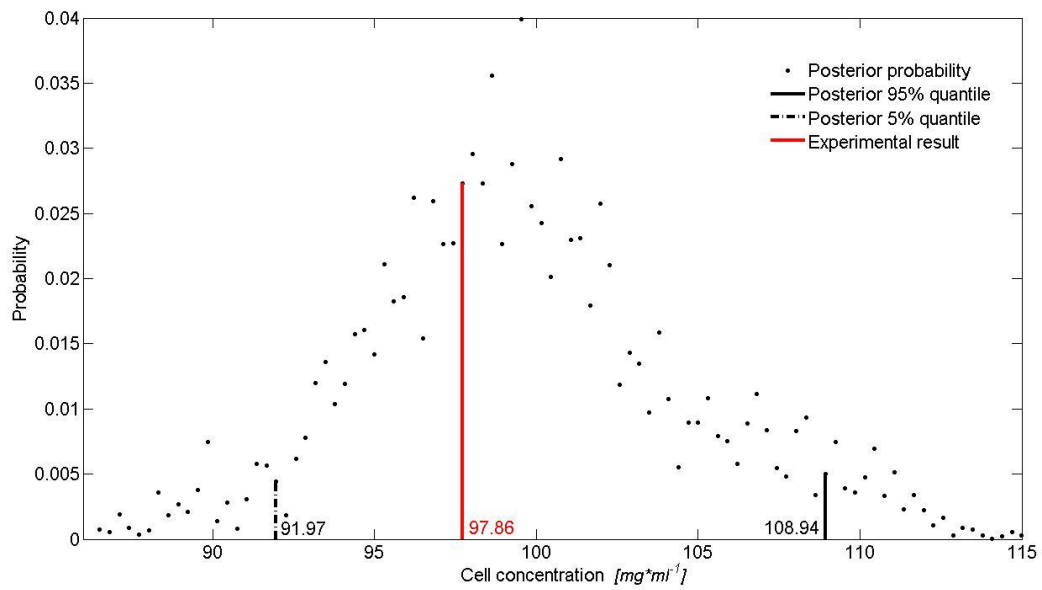
Figure C.27: Posterior probability of cell concentration at time t=5 h Michaelis-Menten-Model
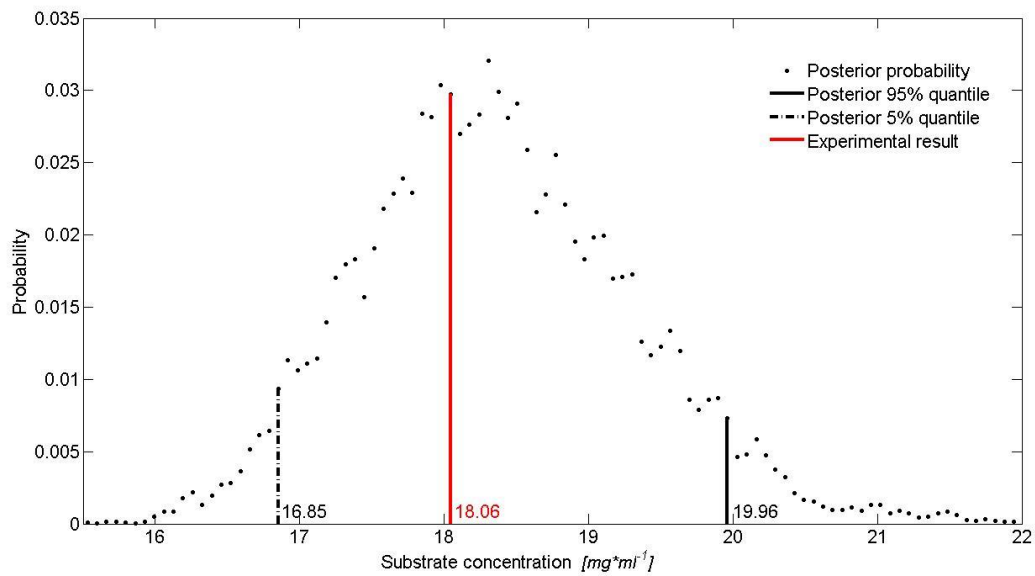
Figure C.28: Posterior probability of substrate concentration at time t=5 h Michaelis-Menten-Model
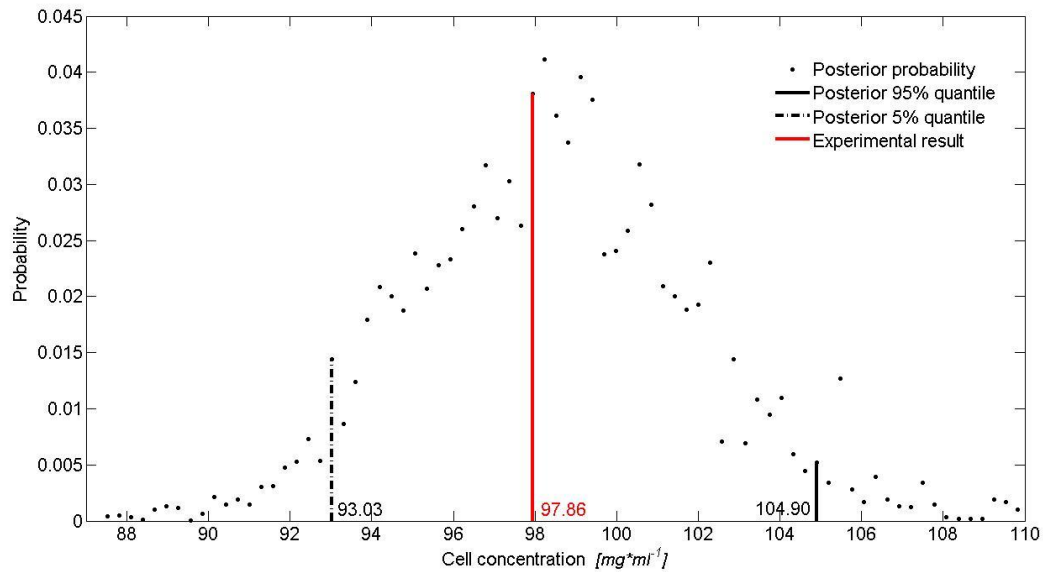
Figure C.29: Posterior probability of cell concentration at time t=5 h Moser-Model
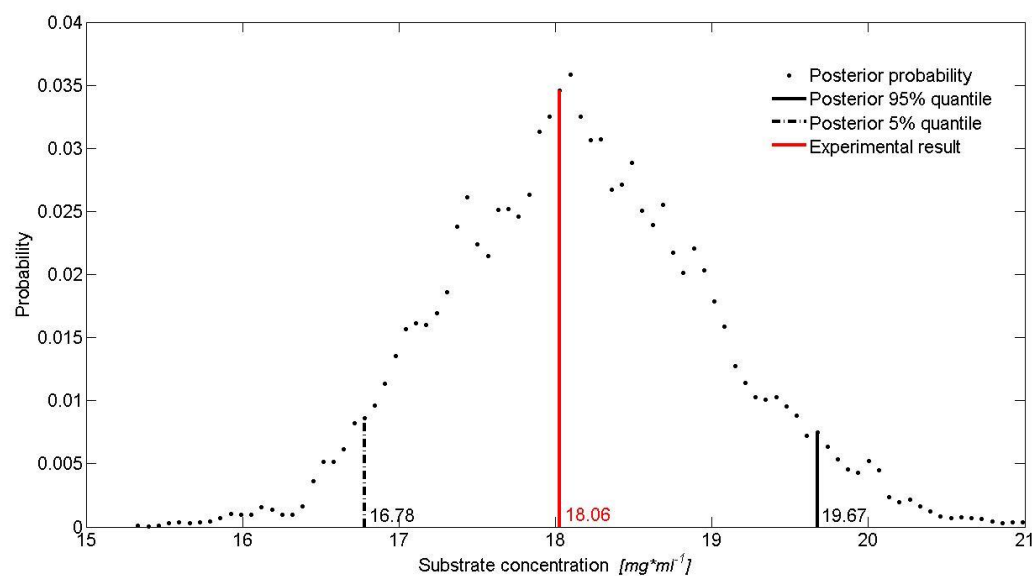
Figure C.30: Posterior probability of substrate concentration at time t=5 h Moser-Model

- Run 2: Application of the posterior mean as candidate vector for model discrimination

  The figures exposed for Run 1 are also valid here as these results are not dependent on the choice of the candidate vector. The results only differ for the relative model probability (compare Table 5).

Annex D – Figures for 5.2.1.2

- Run 1: Application of the parameterization with the highest posterior probability as candidate vector for model discrimination

  o Concergence of Markov-Chains (exemplary)
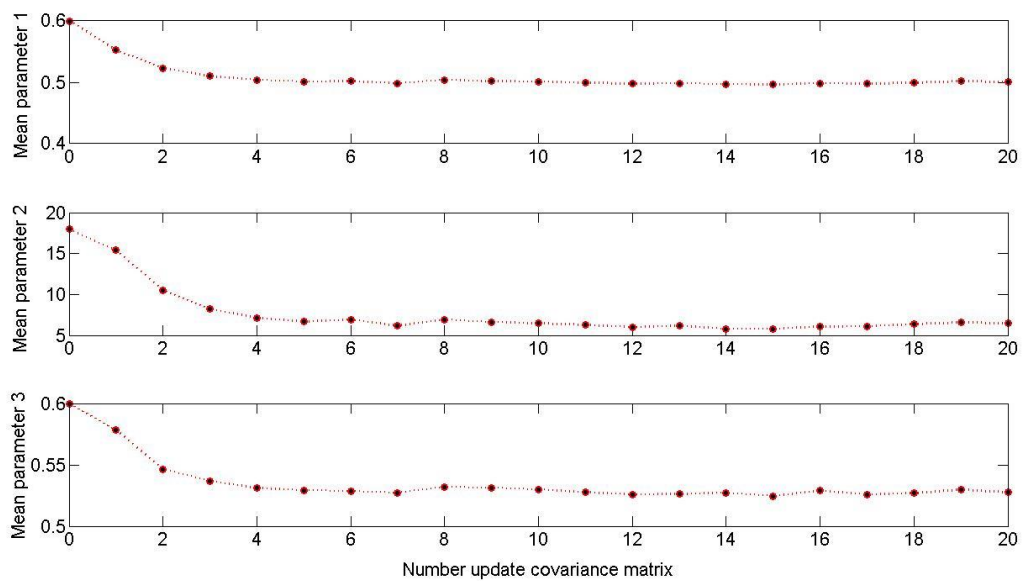
Figure D.31: Convergence mean MM-Model
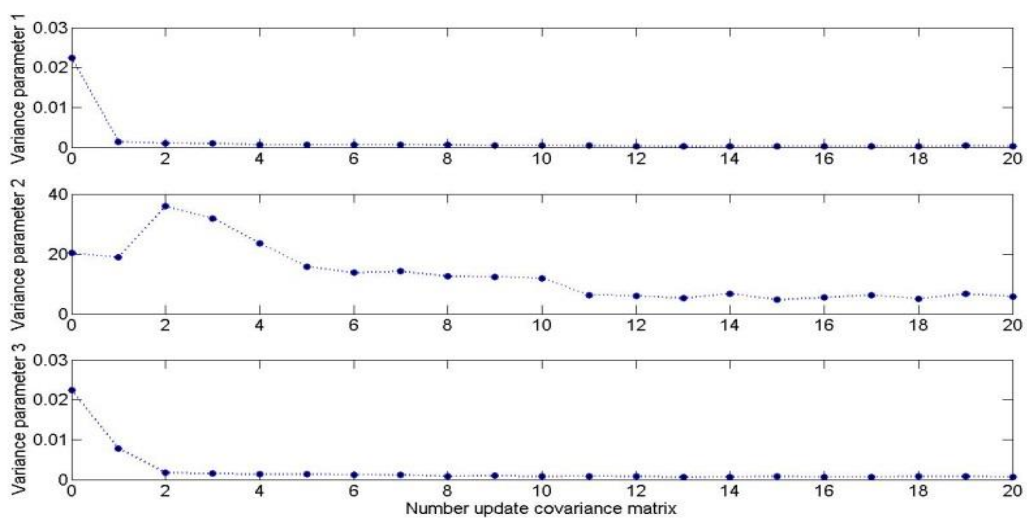
Figure D.32: Convergence variance MM-Model
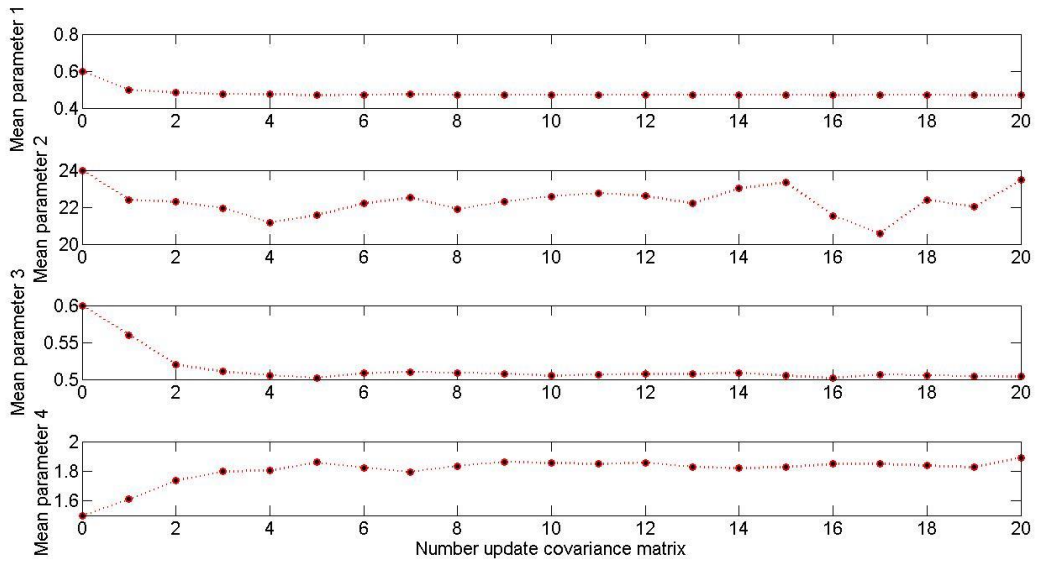
Figure D.33: Convergence mean Moser-Model
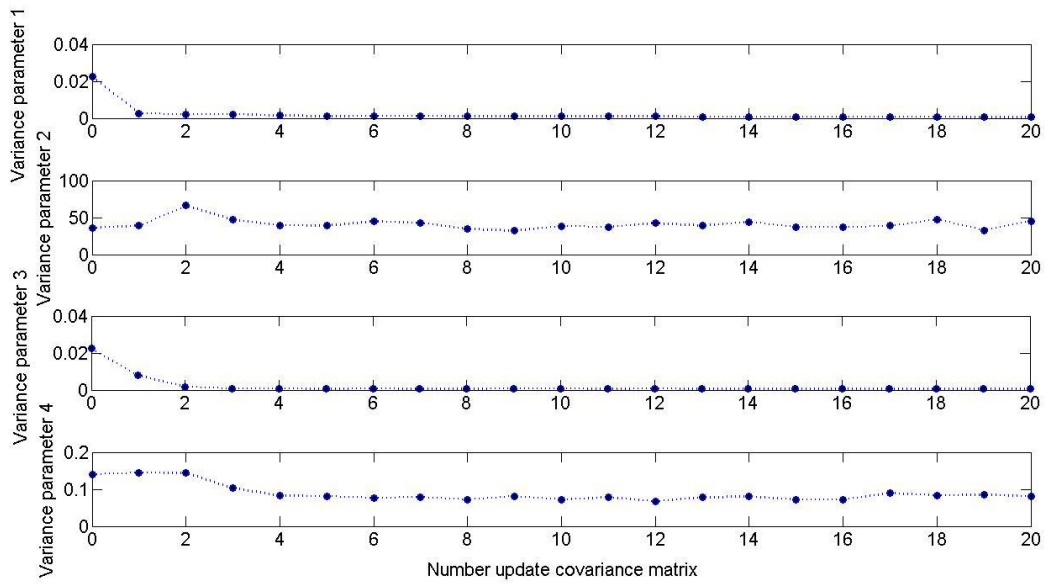
Figure D.34: Convergence variance Moser-Model

156

o Comparisan of Bayesian- and RLS-Model-Fit
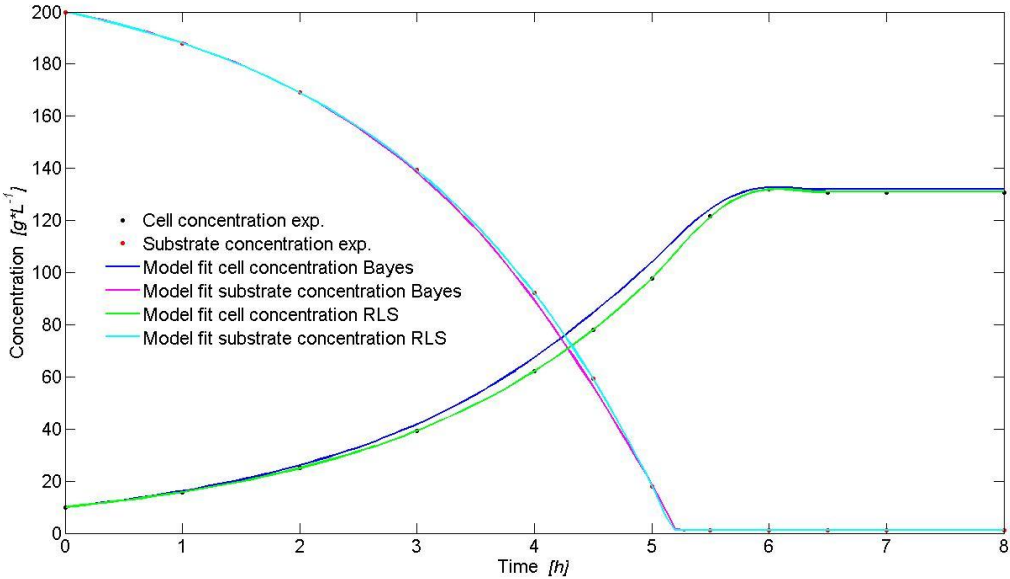
Figure D.35: Bayesian- and RLS-Model-Fit of MM-Model
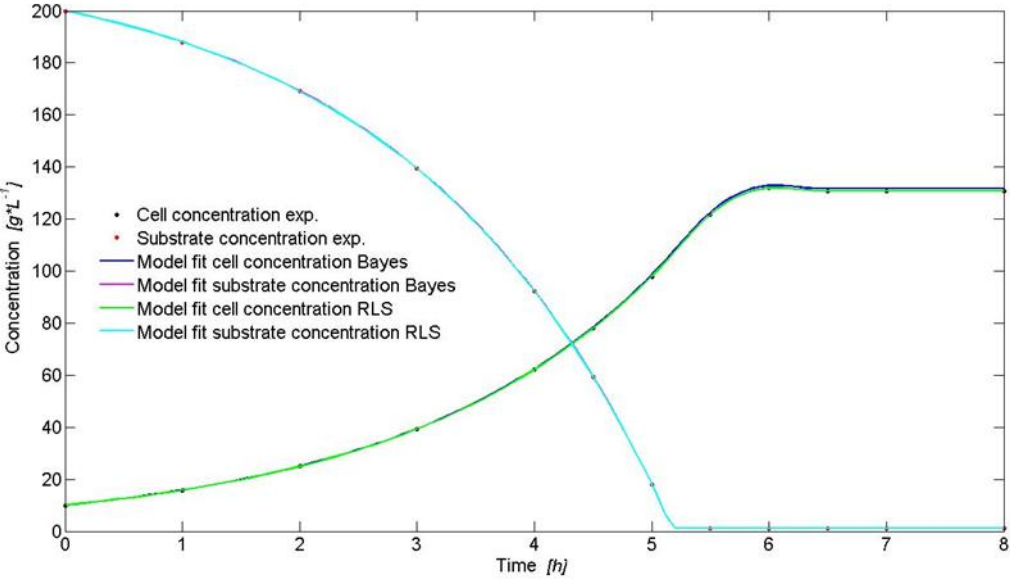
Figure D.36: Bayesian- and RLS-Model-Fit of Moser-Model

157

o   Prior and posterior parameter distributions (exemplary)
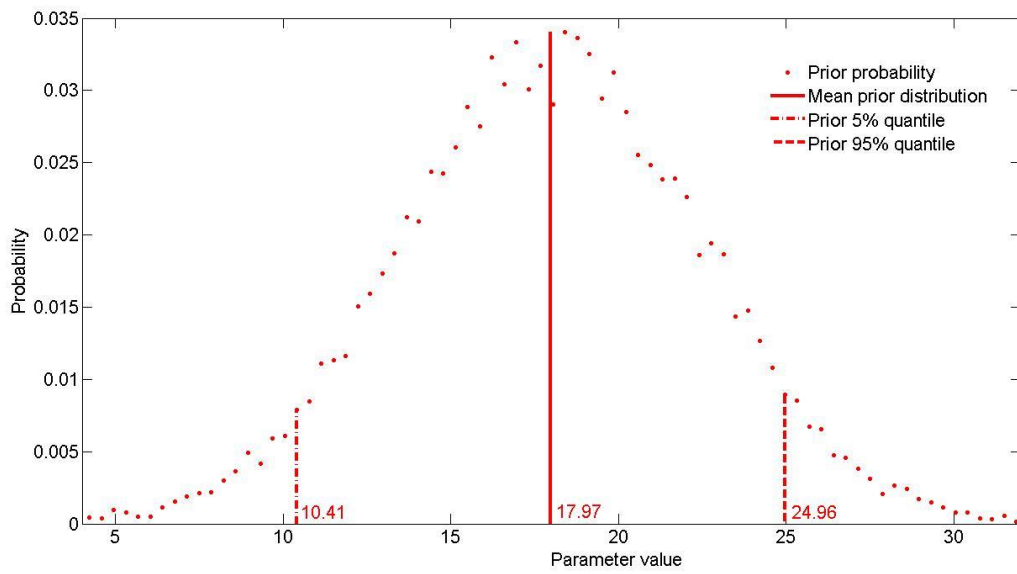
Figure D.37: Prior probability distribution of parameter 2 MM-Model
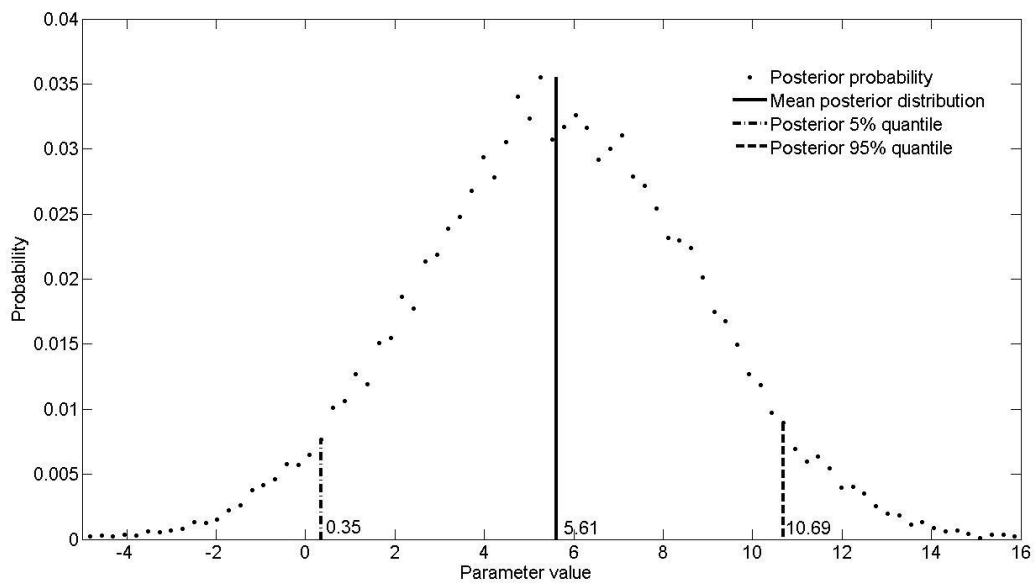
Figure D.38: Posterior probability distribution of parameter 2 MM-Model
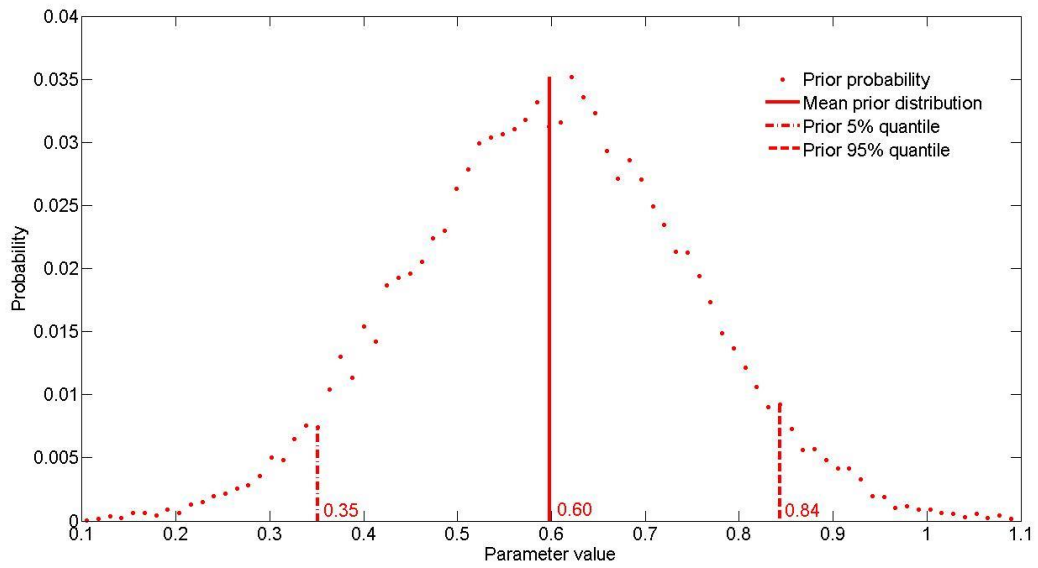
Figure D.39: Prior probability distribution of parameter 3 Moser-Model
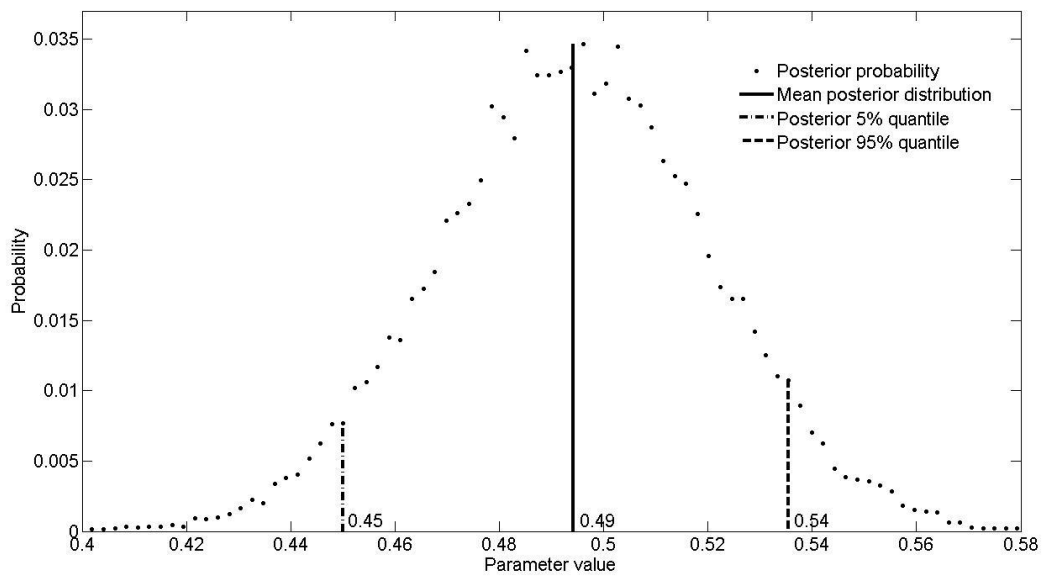
Figure D.40: Posterior probability distribution of parameter 3 Moser-Model
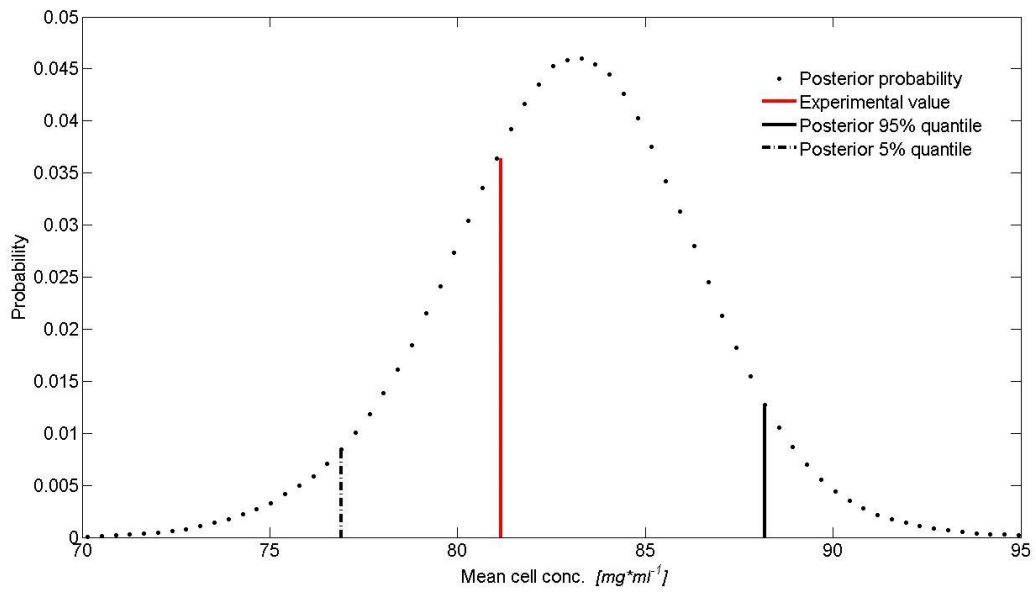
o Model validation

Figure D.41: Validation mean cell concentration MM-Model
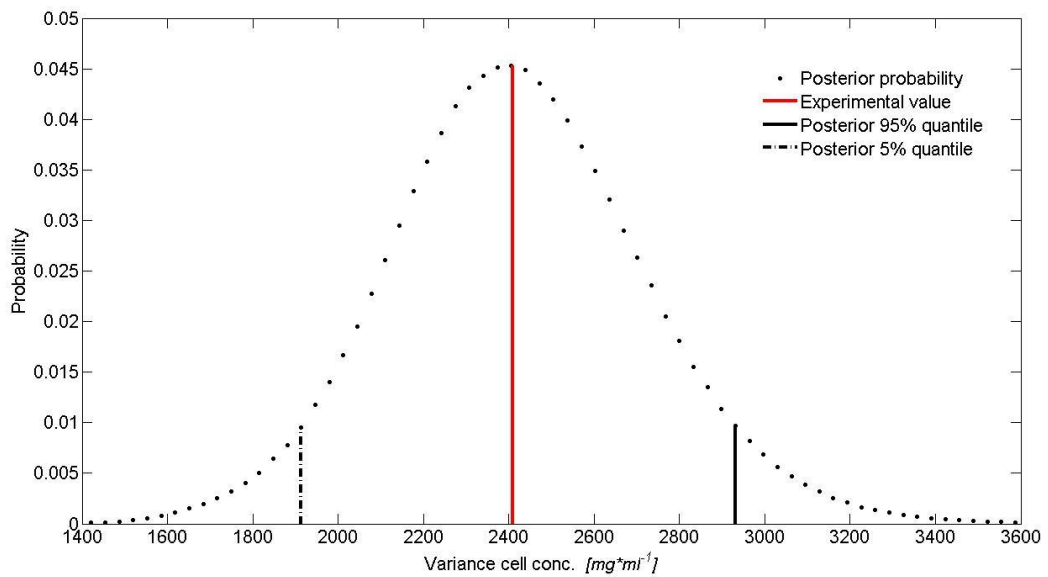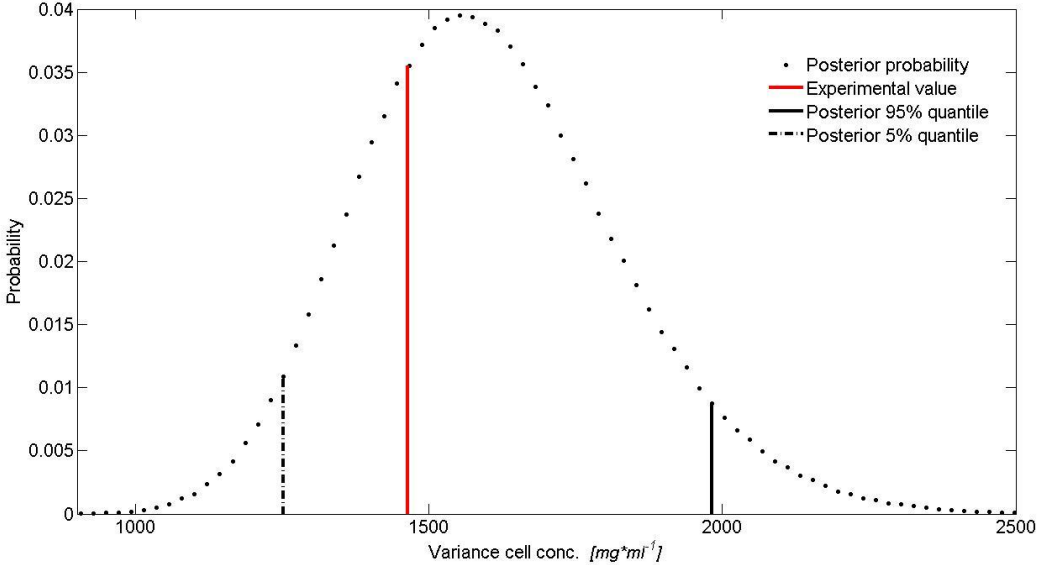

Figure D.42: Validation variance cell concentration MM-Model
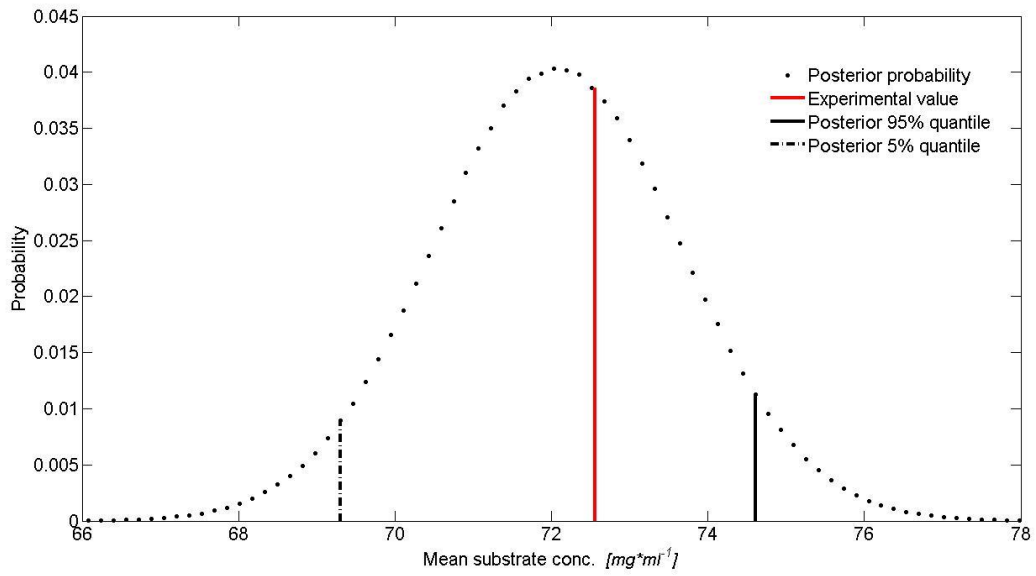
Figure D.43: Validation mean substrate concentration MM-Model
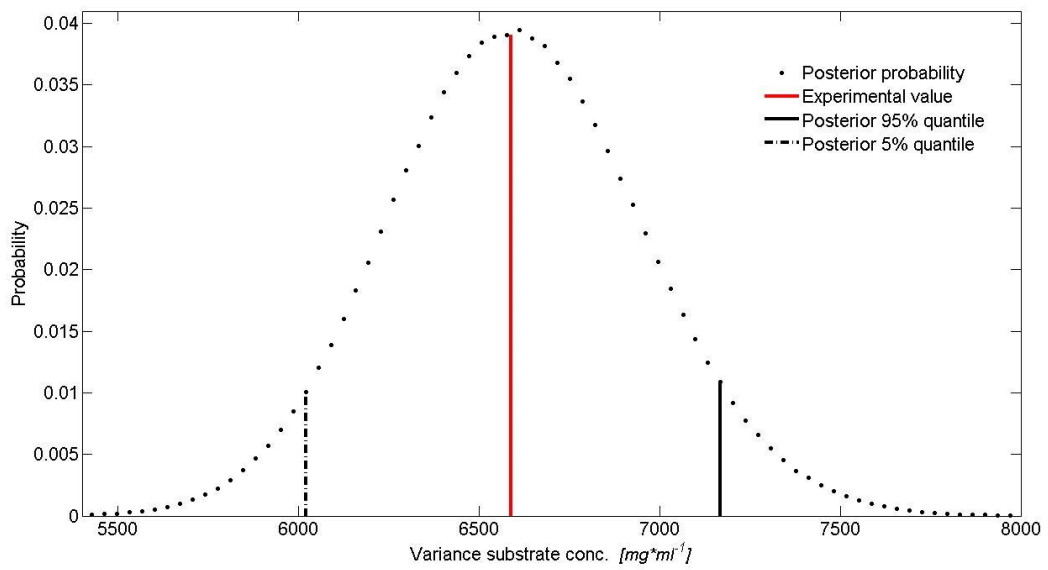
Figure D.44: Validation variance substrate concentration MM-Model
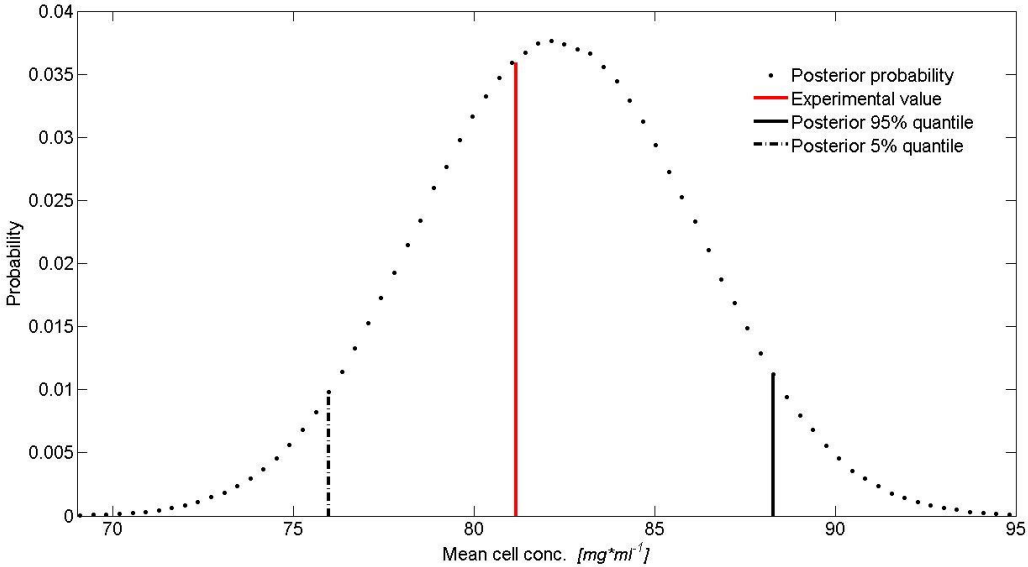
Figure D.45: Validation mean cell concentration Moser-Model
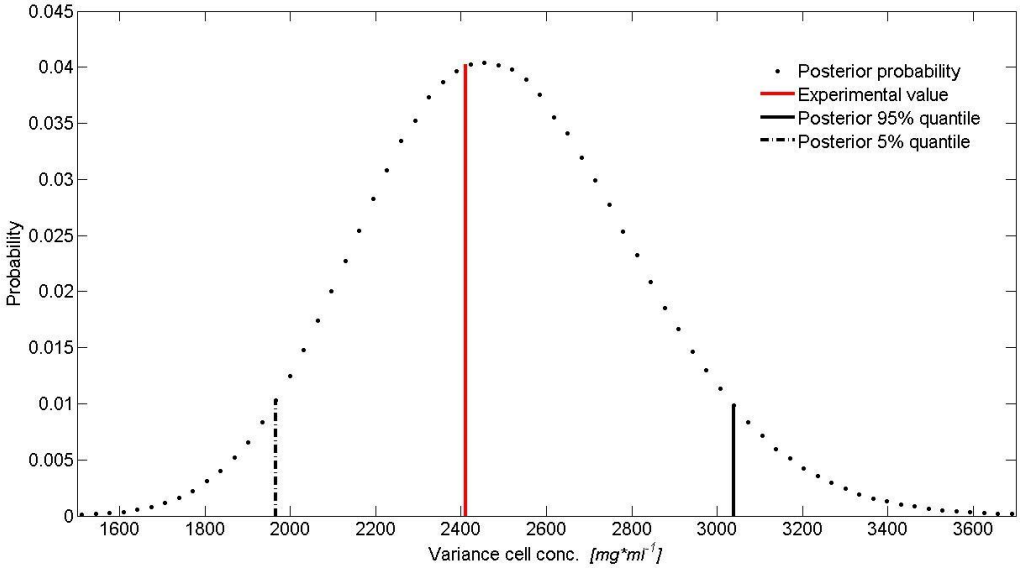
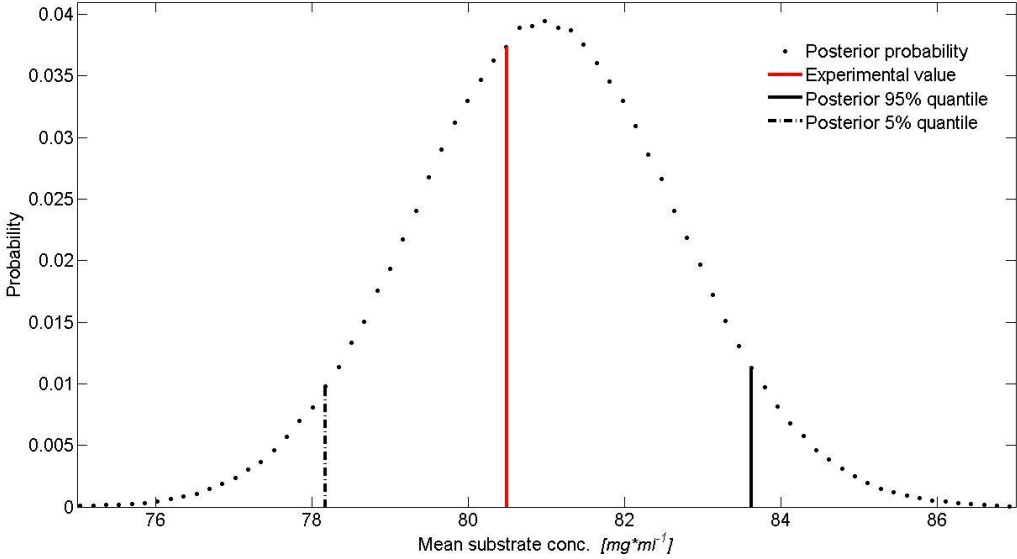Figure D.46: Validation variance cell concentration Moser-Model
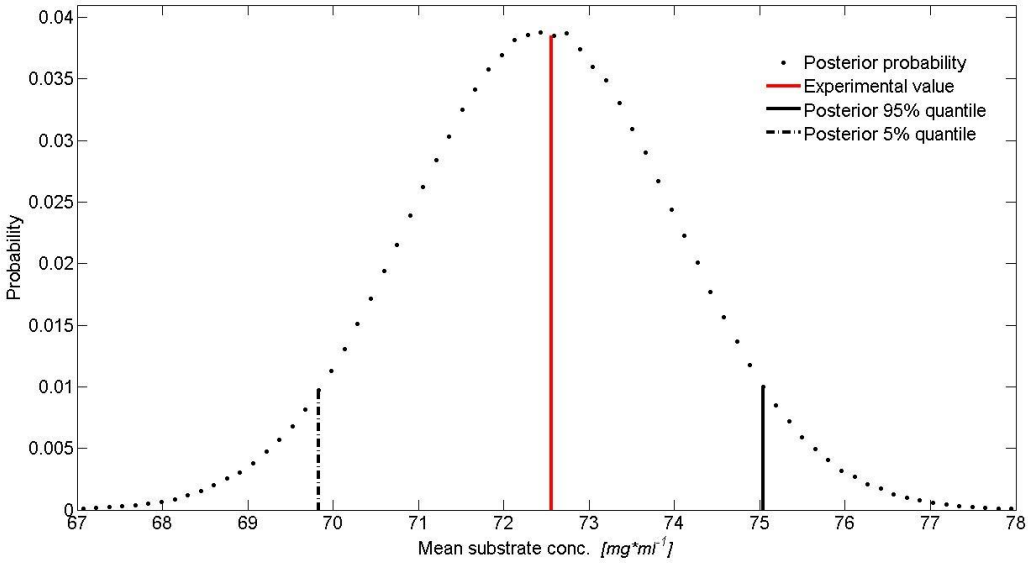
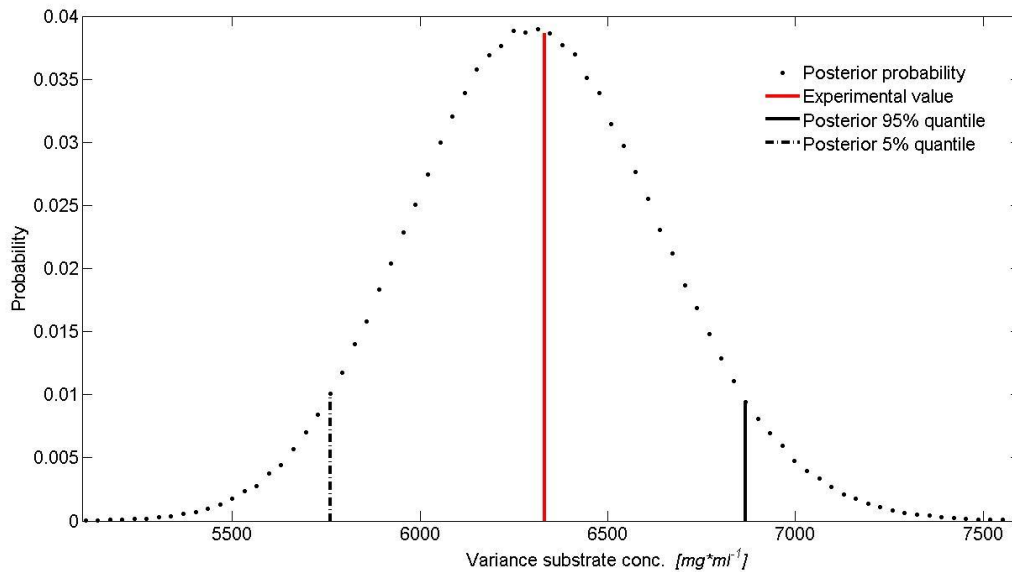Figure D.47: Validation mean substrate concentration Moser-Model
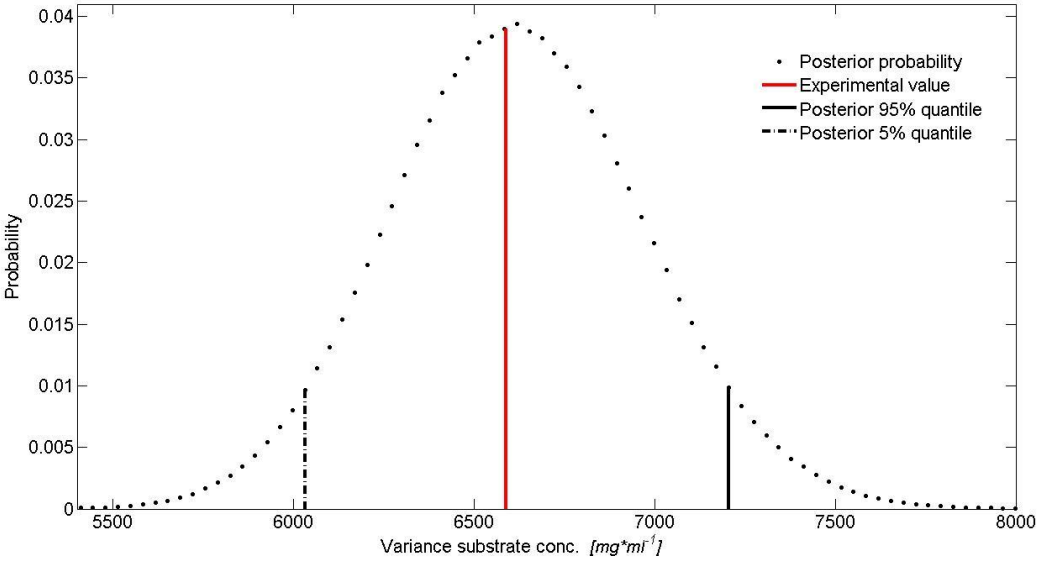
Figure D.48: Validation variance substrate concentration Moser-Model

o   Posterior probability distribution of experimental observations (exemplary)
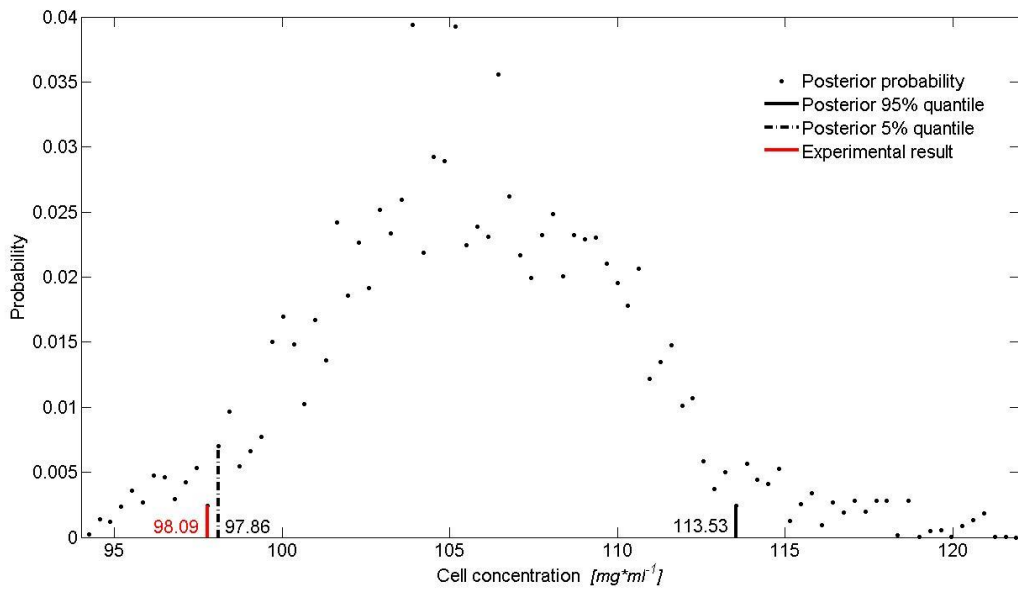
Figure D.49: Posterior probability of cell concentration at time t=5 h Michaeli-Menten-Model
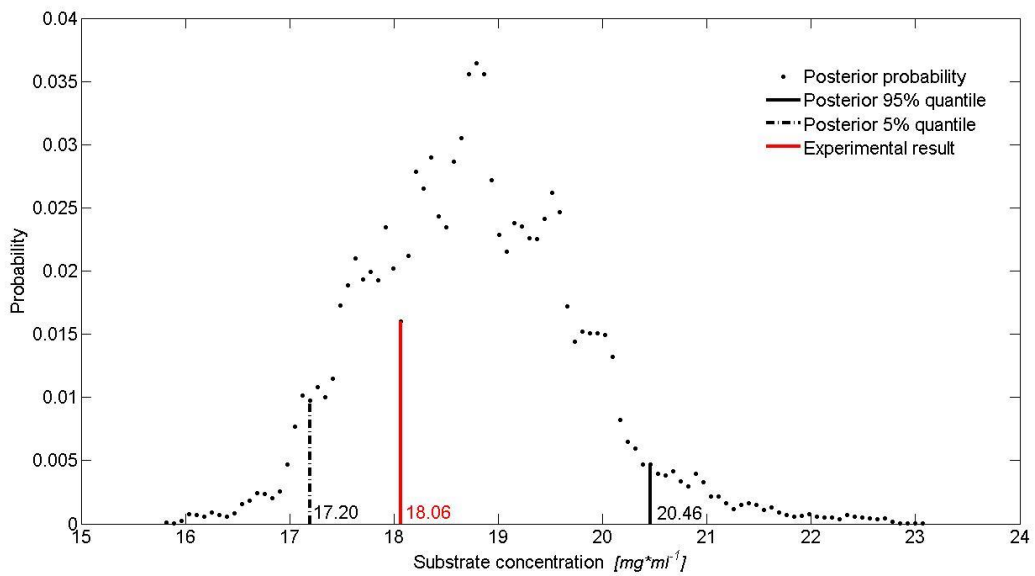
Figure D.50: Posterior probability of substrate concentration at time t=5 h Michaeli-Menten-Model
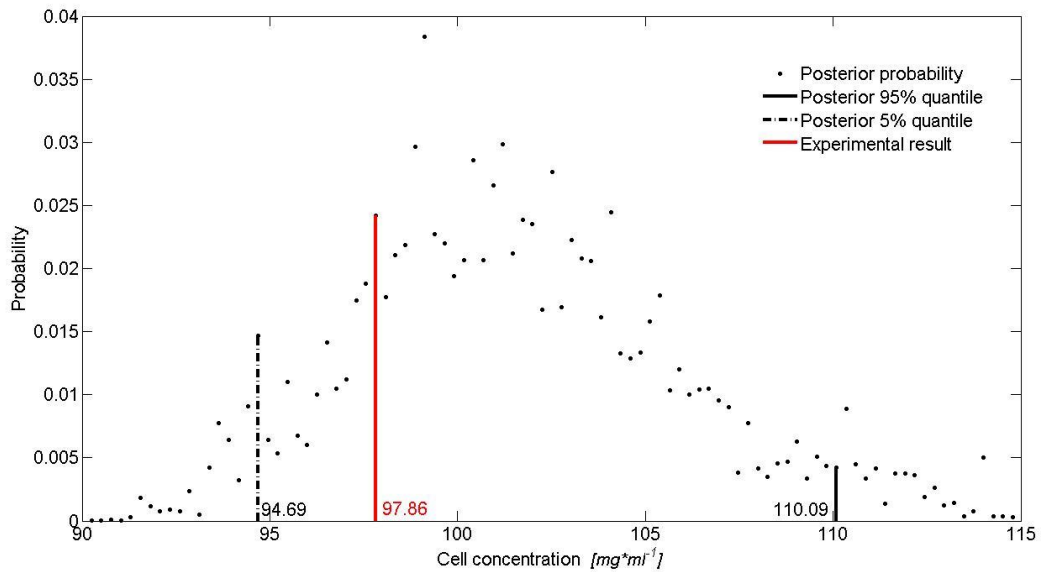
Figure D.51: Posterior probability of cell concentration at time t=5 h Moser-Model
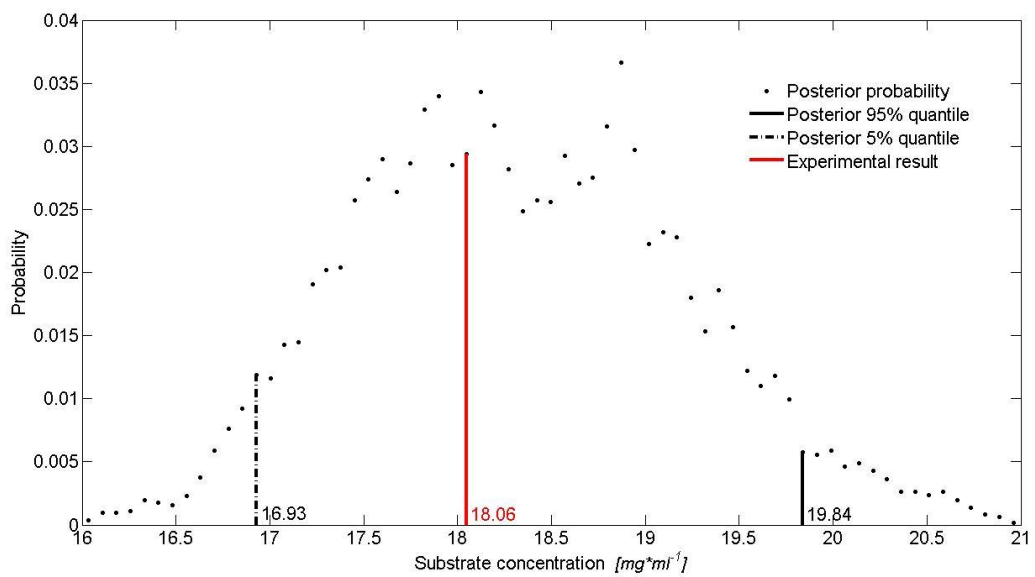
Figure D.52: Posterior probability of substrate concentration at time t=5 h Moser-Model

- Run 2: Application of the posterior mean as candidate vector for model discrimination

  The figures exposed for Run 1 are also valid here as these results are not dependent on the choice of the candidate vector. The results only differ for the relative model probability (compare Table 5.19).

165

Annex E – Figures for 5.2.1.3

Application of the parameterization with the highest posterior probability as candidate vector for model discrimination

- o Concergence of Markov-Chains (exemplary)
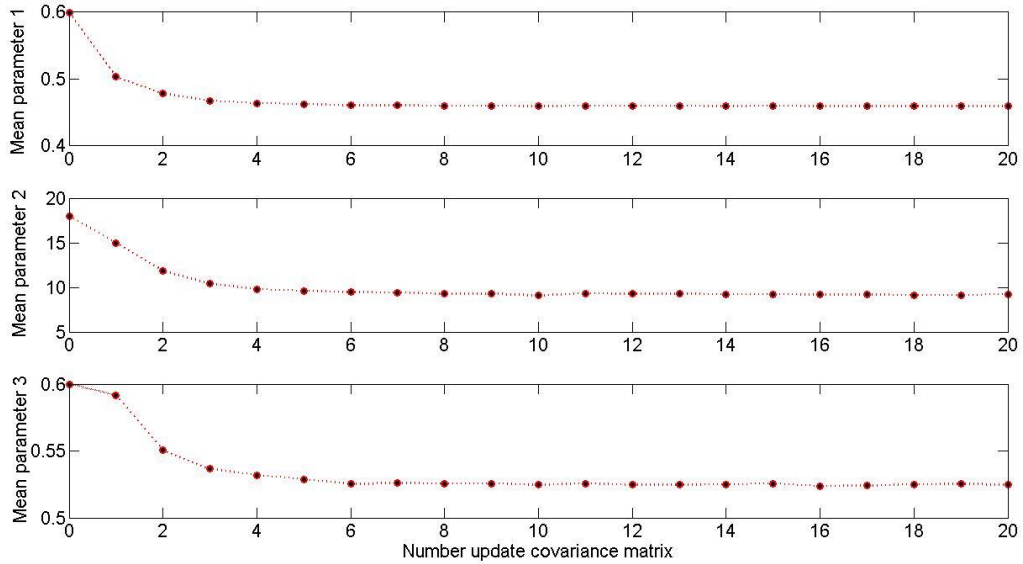
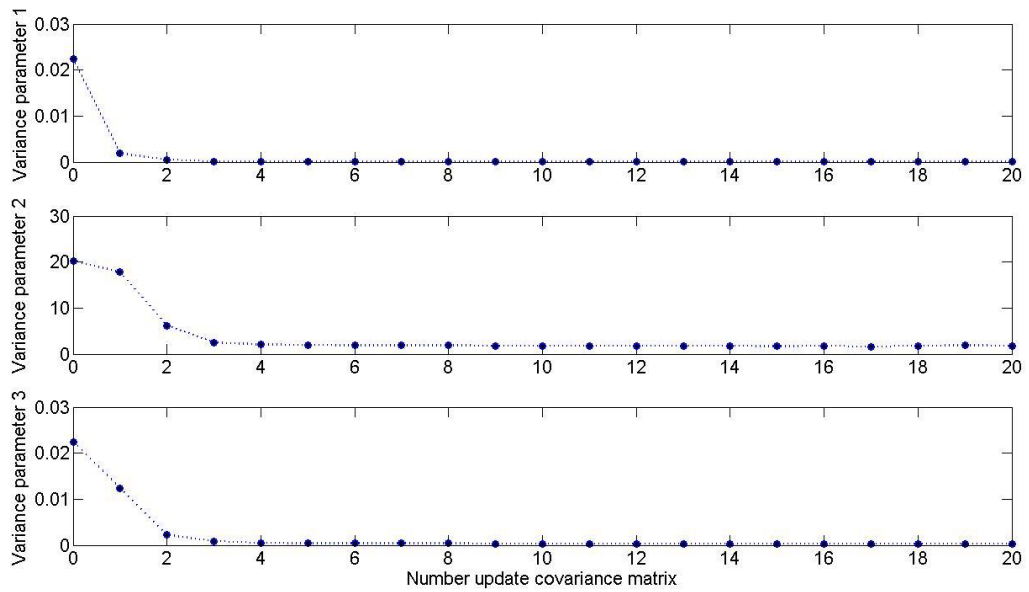
Figure E.53: Convergence mean MM-Model
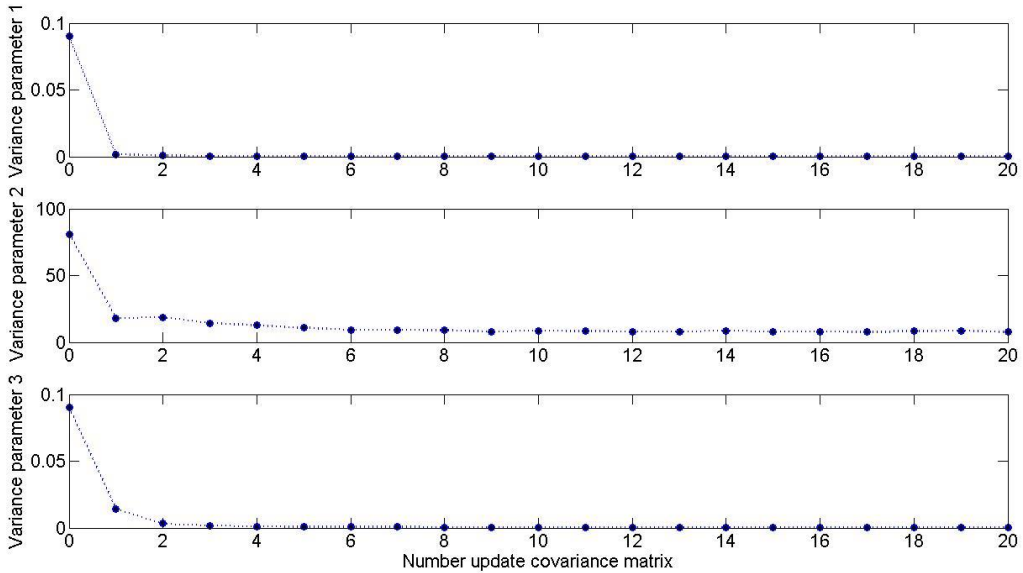
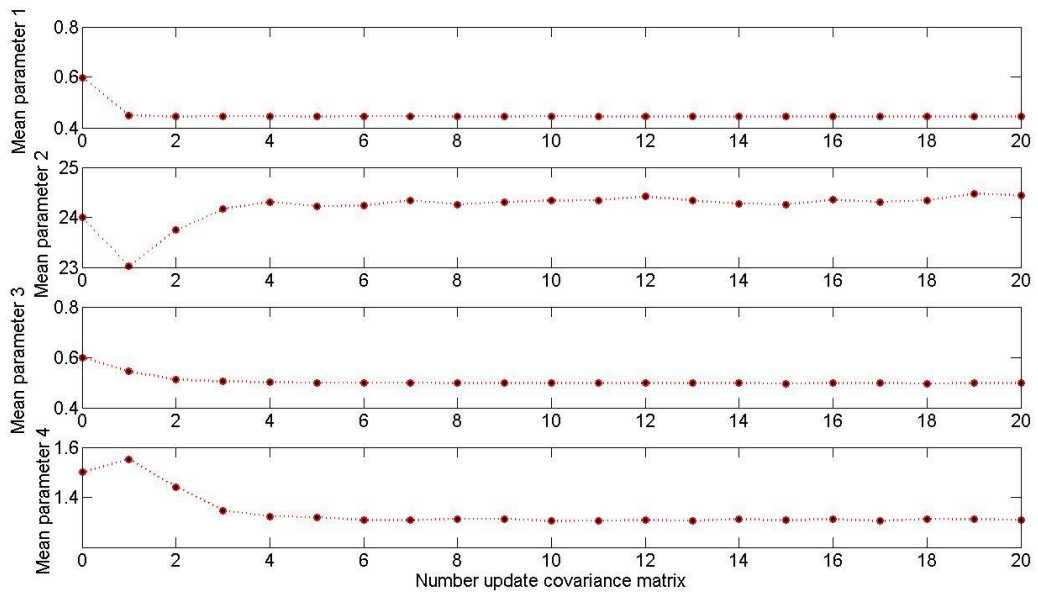Figure E.54: Convergence variance MM-Model

166
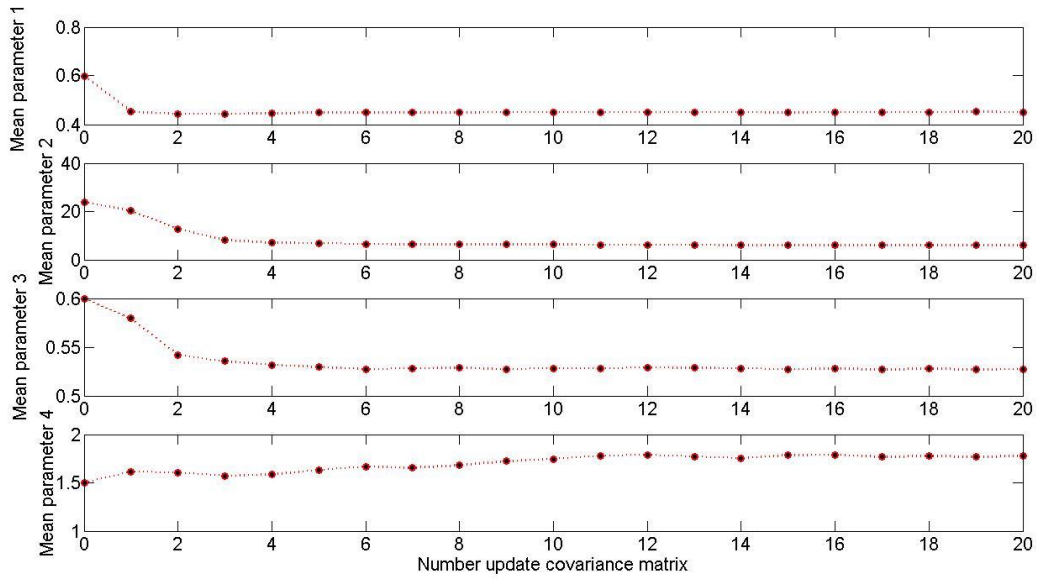
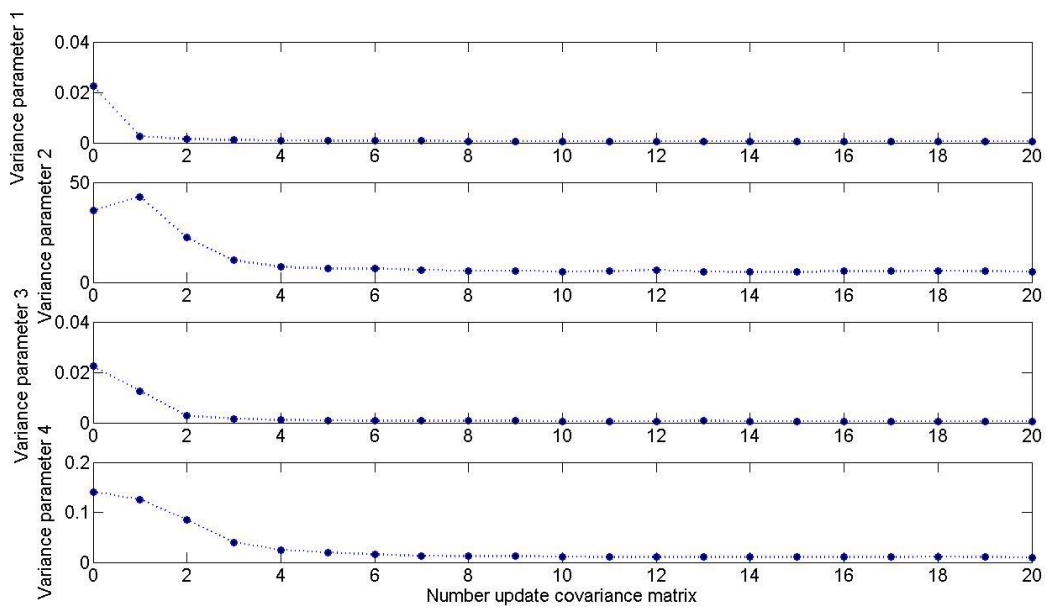Figure E.55: Convergence mean Moser-Model
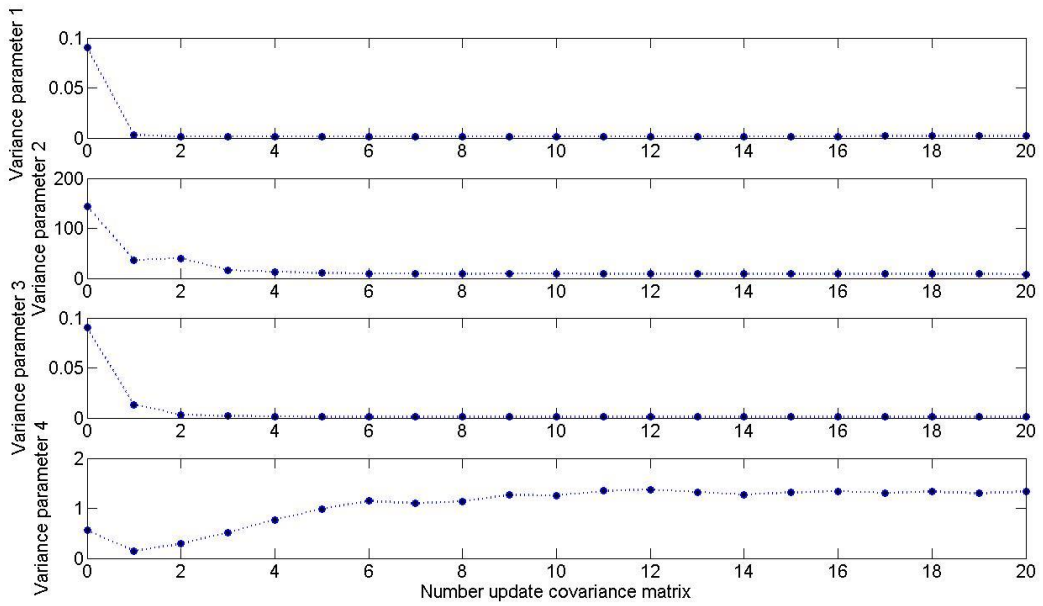
Figure E.56: Convergence variance Moser-Model

o Comparisan of Bayesian- and RLS-Model-Fit
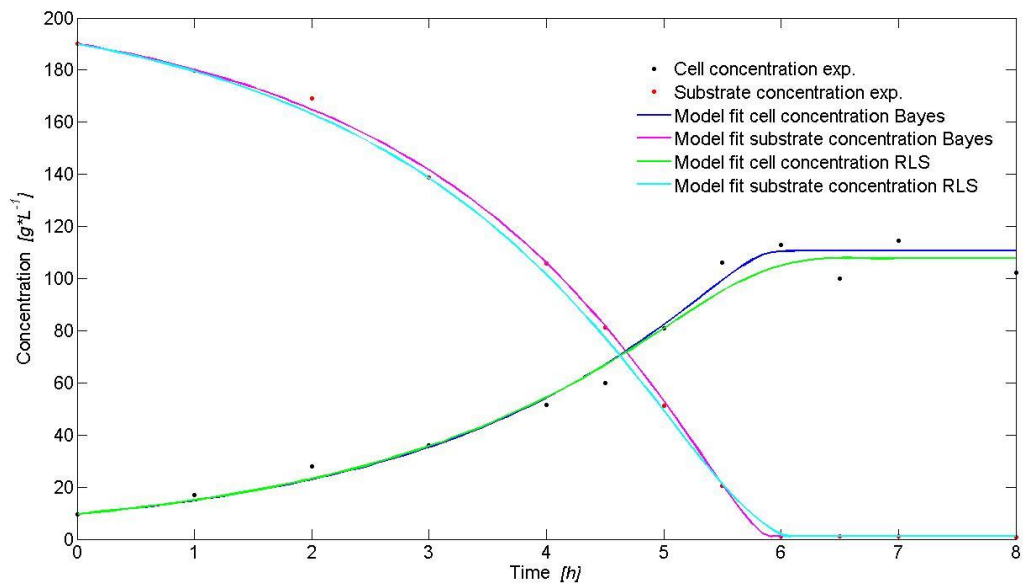
Figure E.57: Bayesian- and RLS-Model-Fit of MM-Model
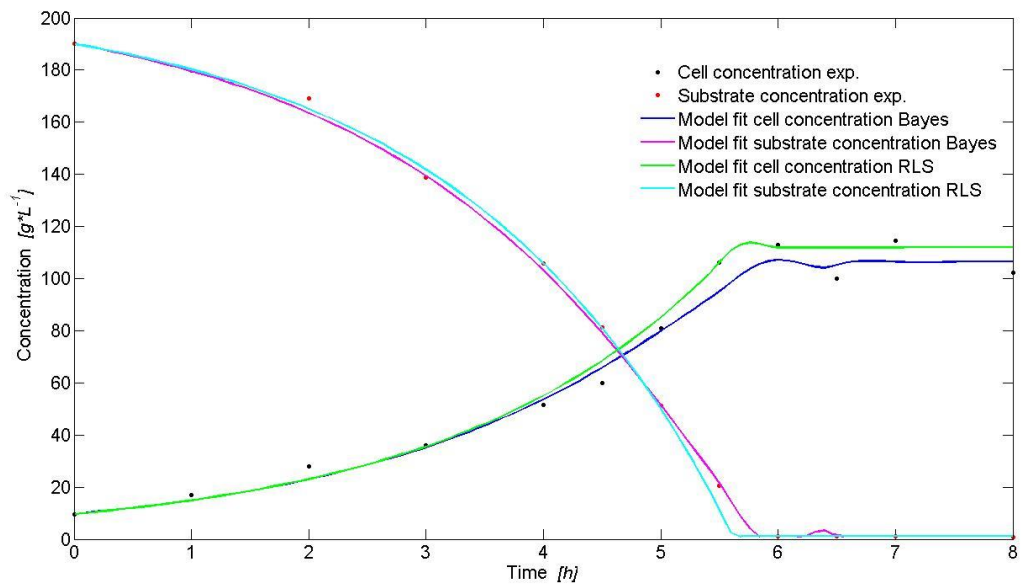
Figure E.58: Bayesian- and RLS-Model-Fit of Moser-Model

o   Prior and posterior parameter distributions (exemplary)
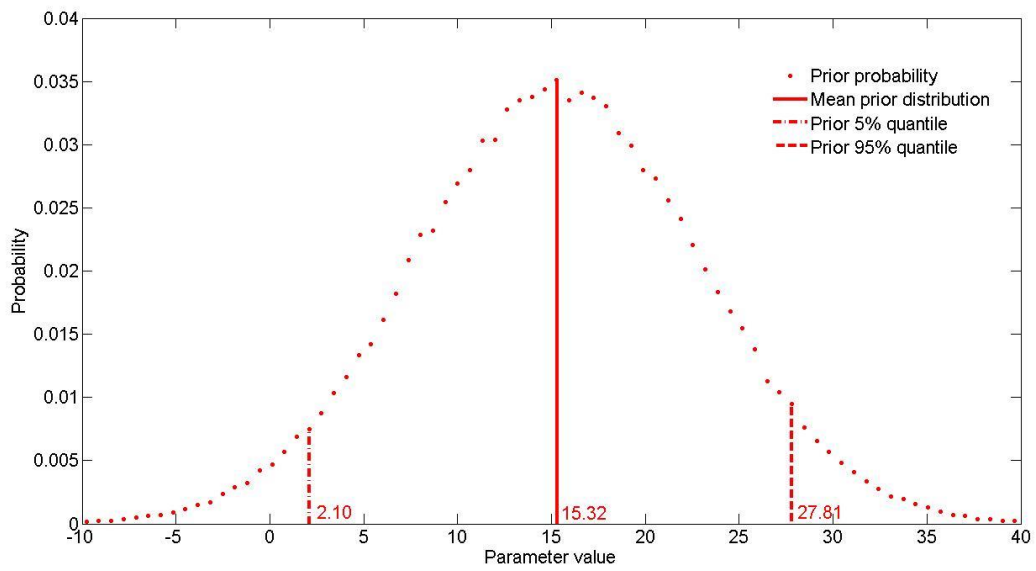
Figure E.59: Prior proabability distribution of parameter 2 MM-Model
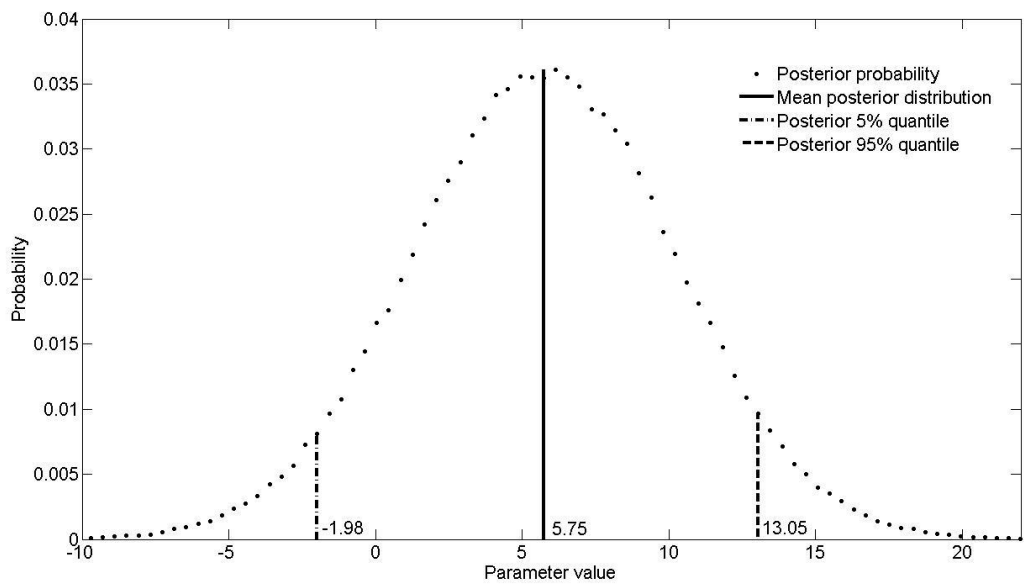
Figure E.60: Posterior proabability distribution of parameter 2 MM-Model
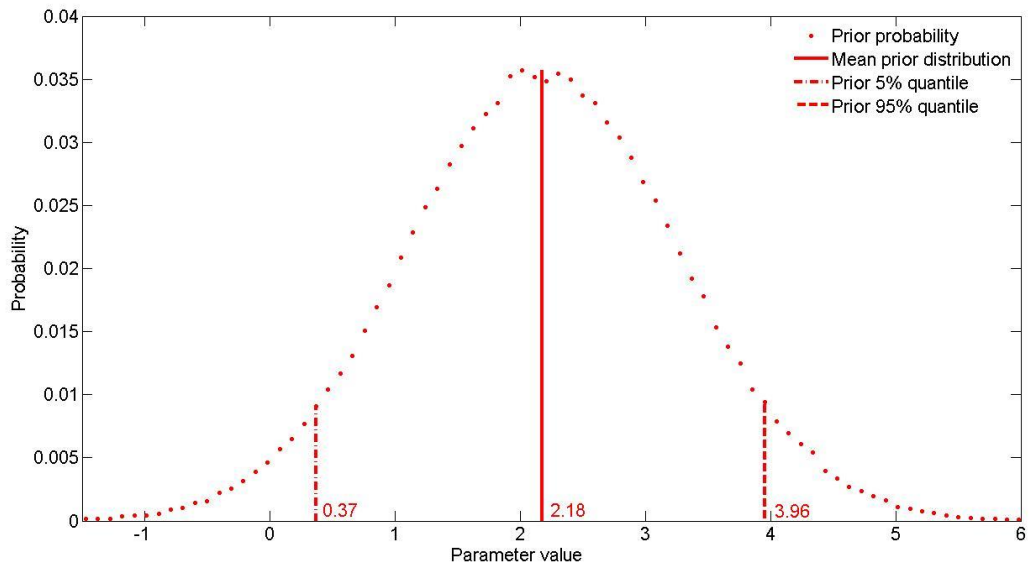
Figure E.61: Prior proability distribution of parameter 3 Moser-Model
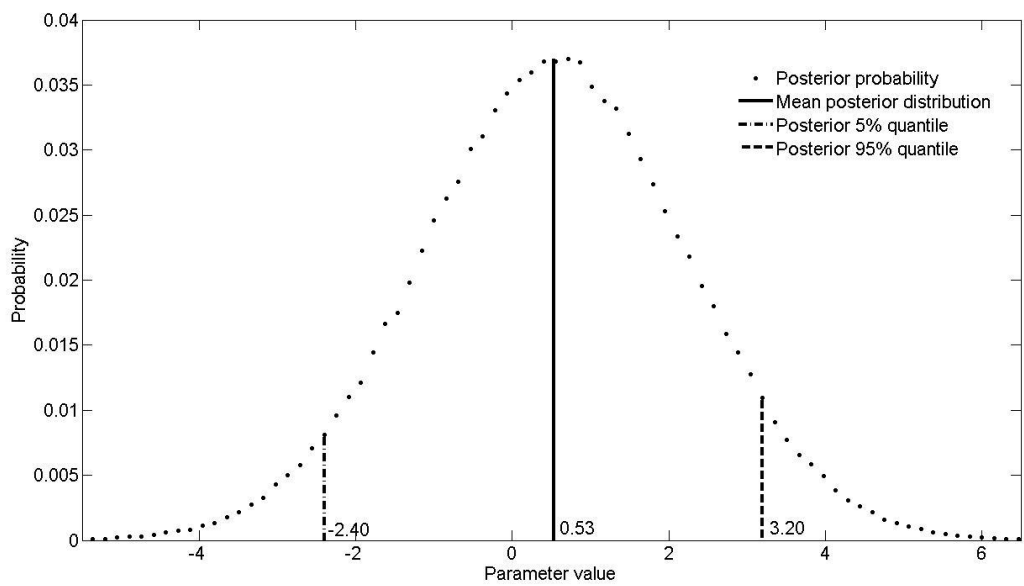
Figure E.62: Posterior proability distribution of parameter 3 Moser-Model
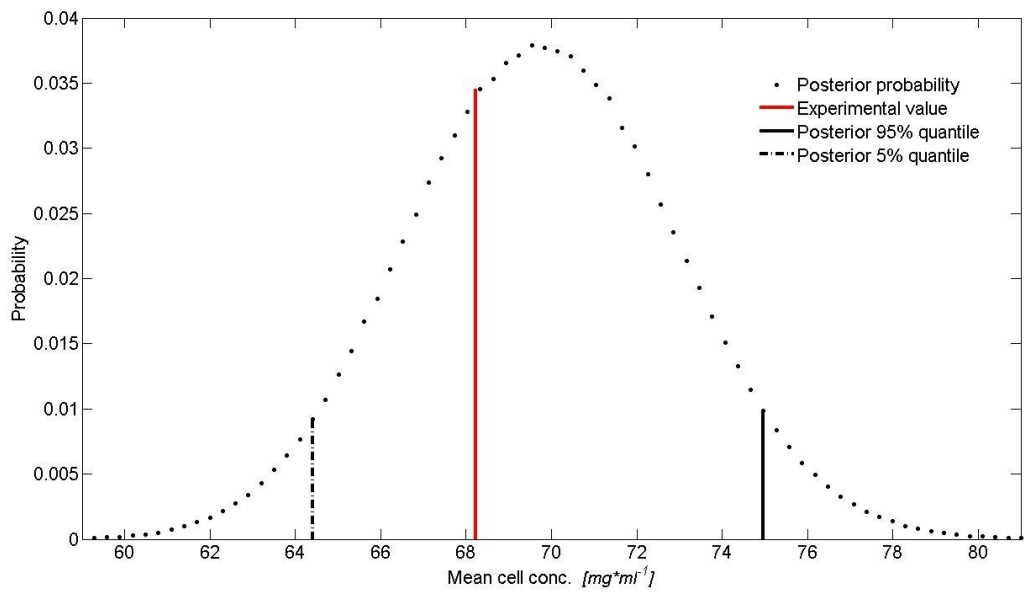
o Model validation

Figure E.63: Validation mean cell concentration MM-Model
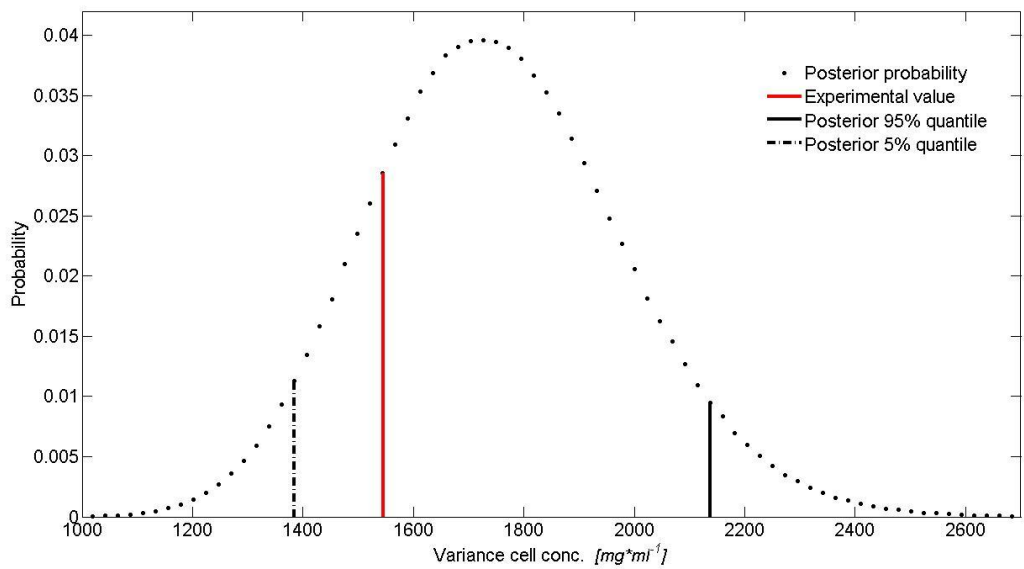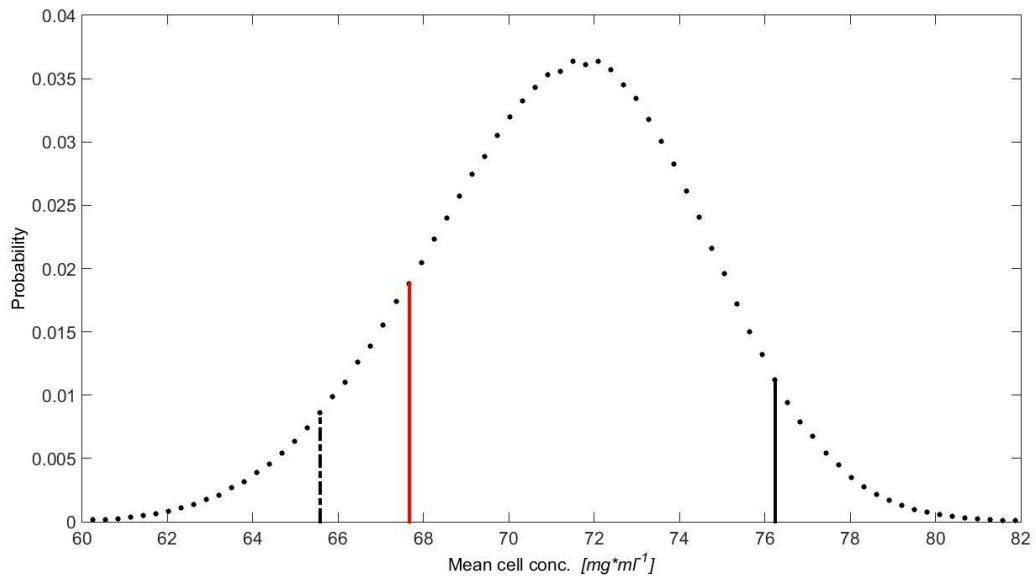

Figure E.64: Validation variance cell concentration MM-Model
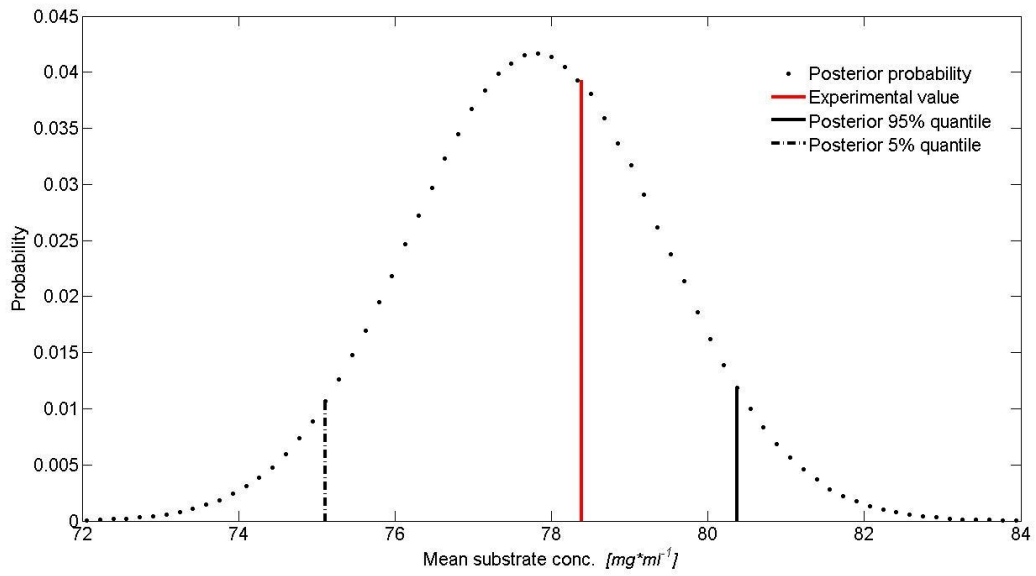
Figure E.65: Validation mean substrate concentration MM-Model
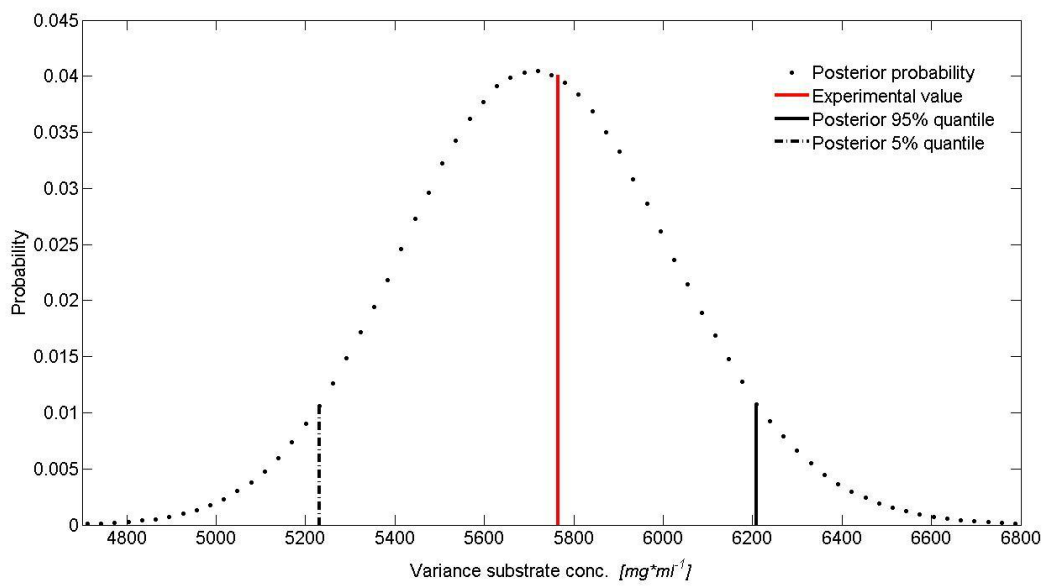
Figure E.66: Validation variance substrate concentration MM-Model
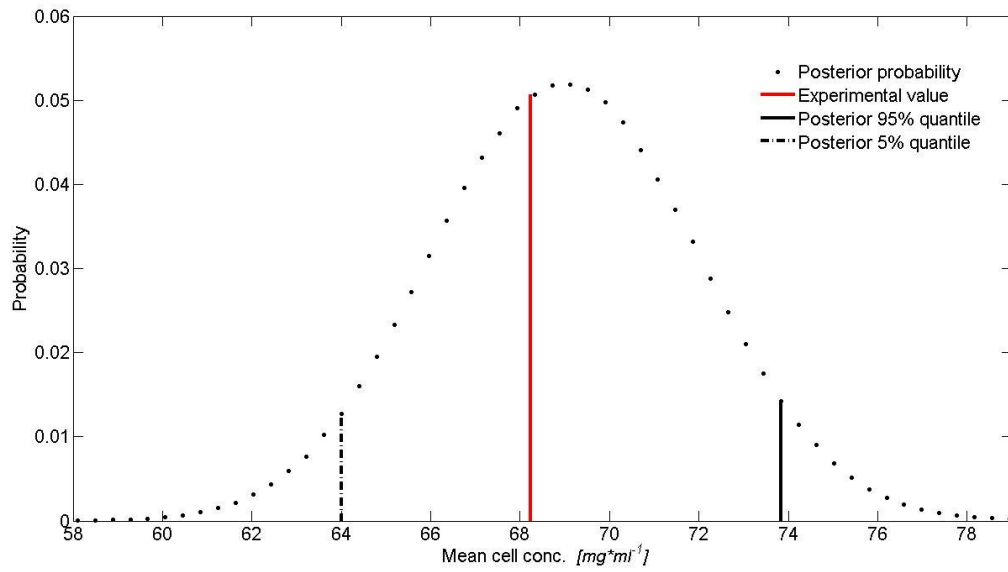
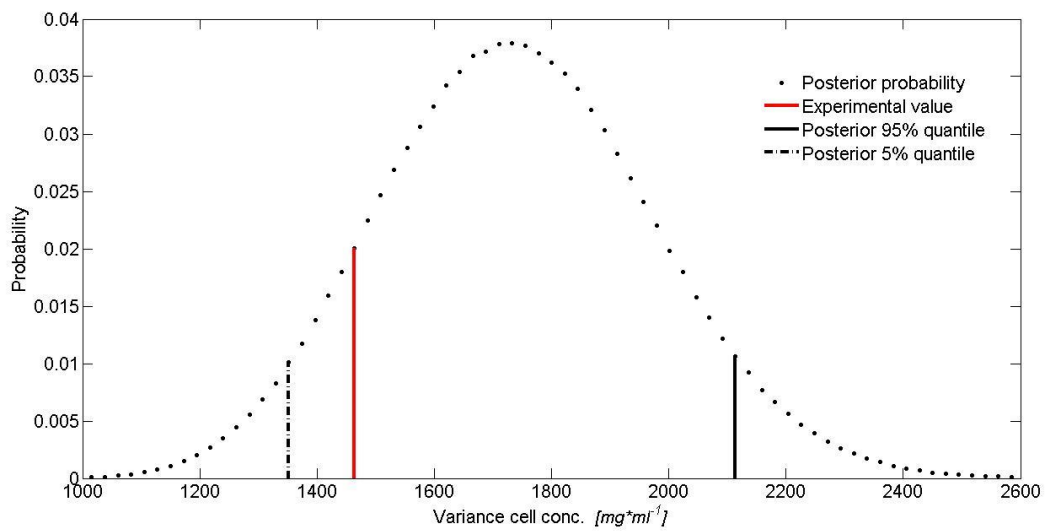Figure E.67: Validation mean cell concentration Moser-Model
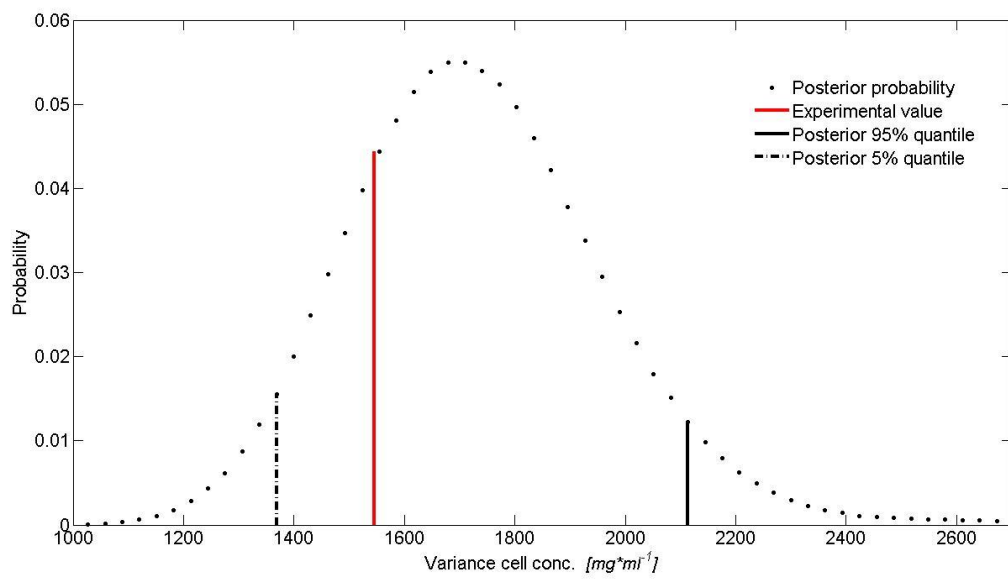
Figure E.68: Validation variance cell concentration Moser-Model

173
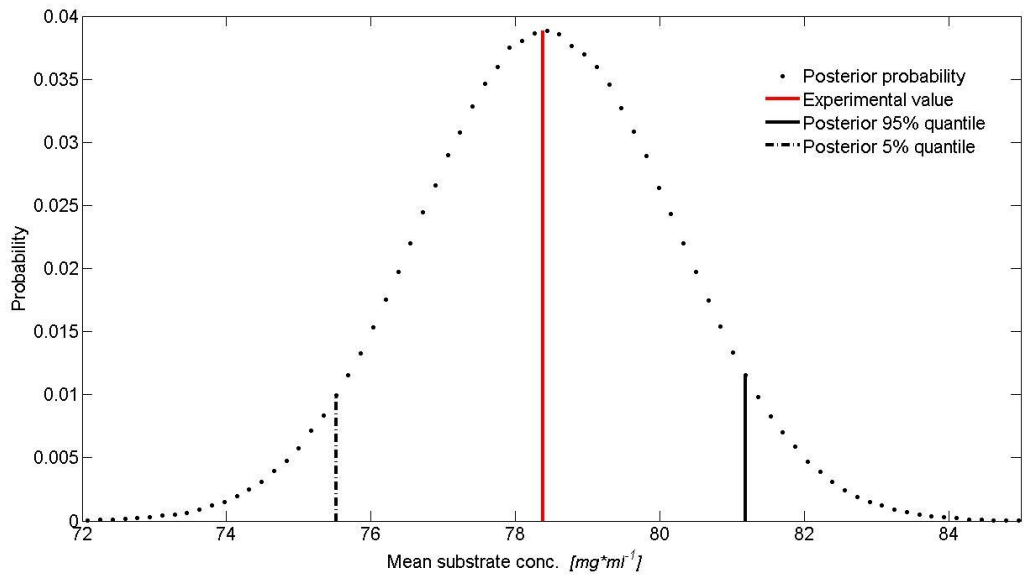
Figure E.69: Validation mean substrate concentration Moser-Model
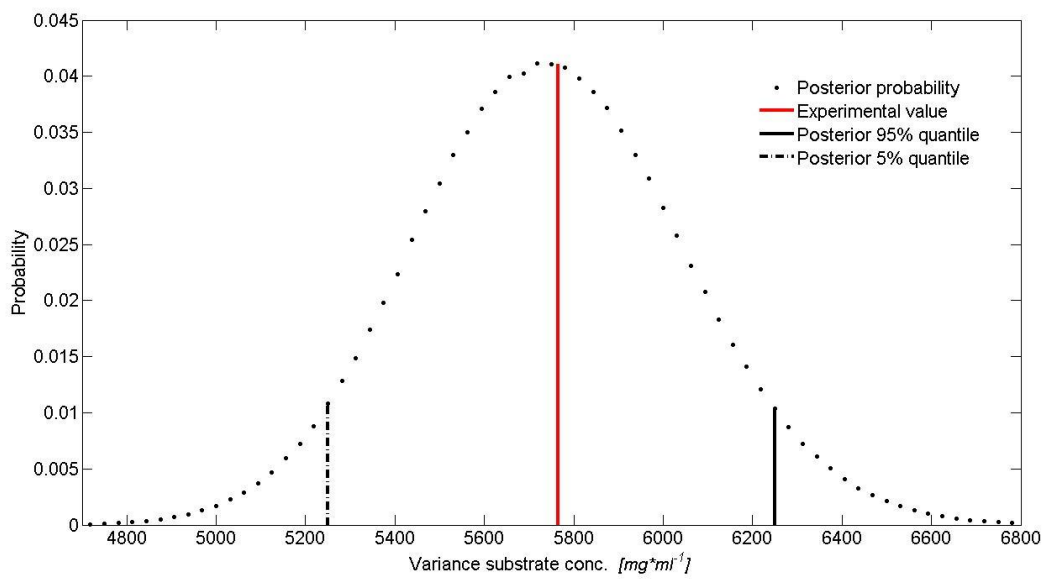
Figure E.70: Validation variance substrate concentration Moser-Model

o   Posterior probability distribution of experimental observations (exemplary)
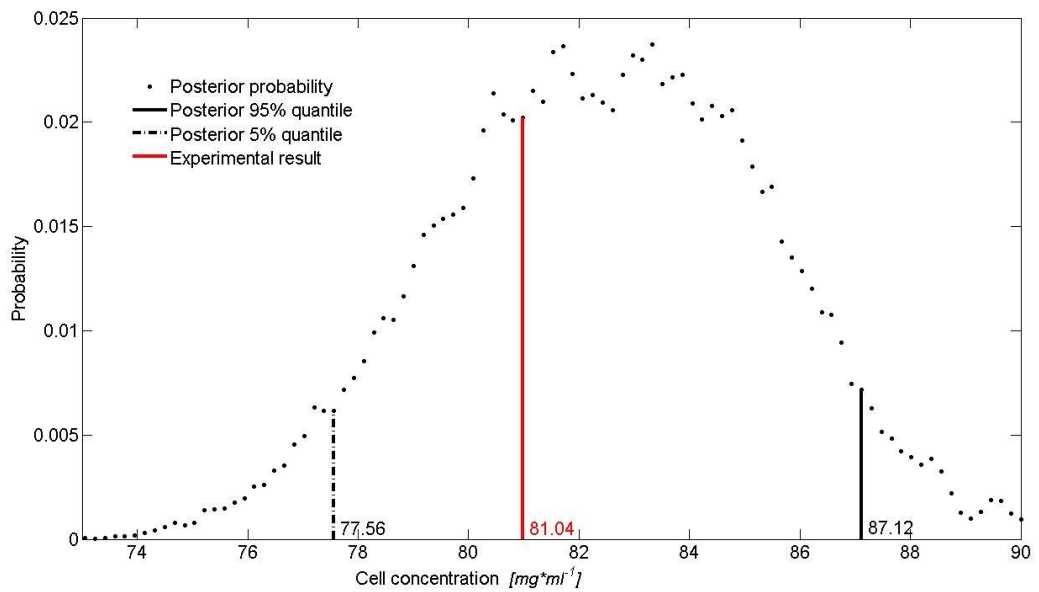
Figure E.71: Posterior probability of cell concentration at time t=5 h Michaeli-Menten-Model
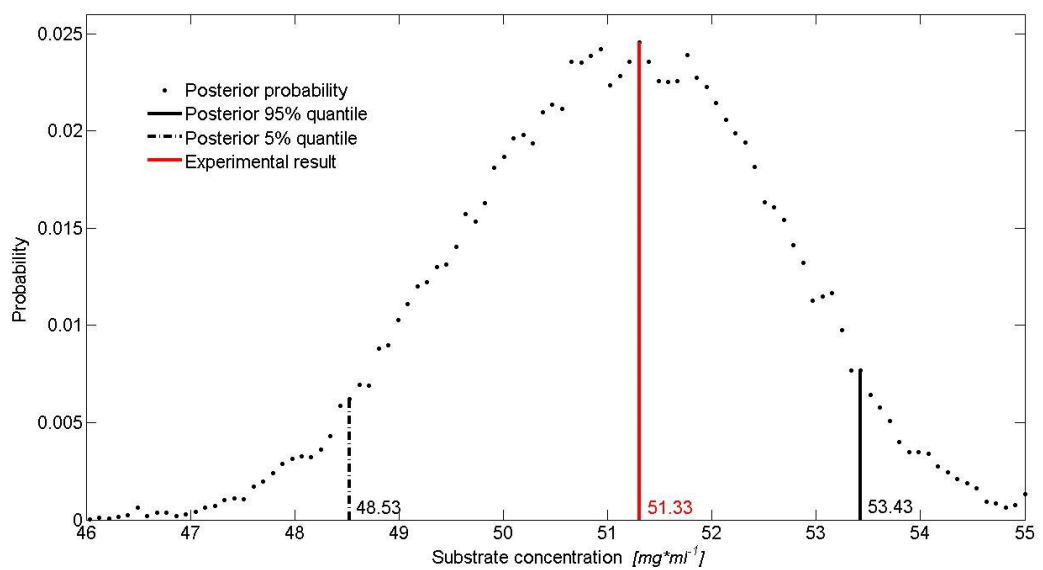
Figure E.72: Posterior probability of substrate concentration at time t=5 h Michaeli-Menten-Model
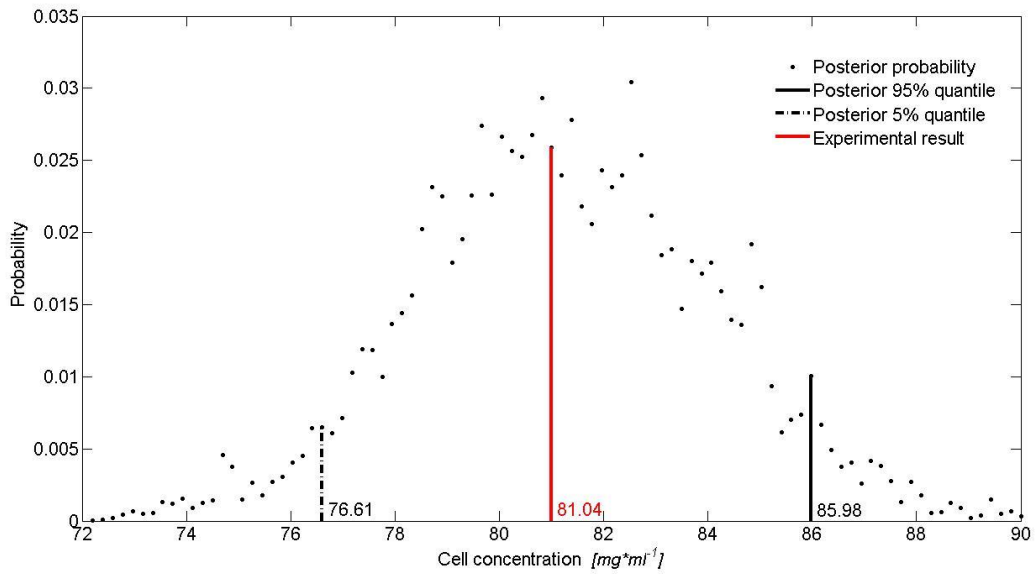
Figure E.73: Posterior probability of cell concentration at time t=5 h Moser-Model
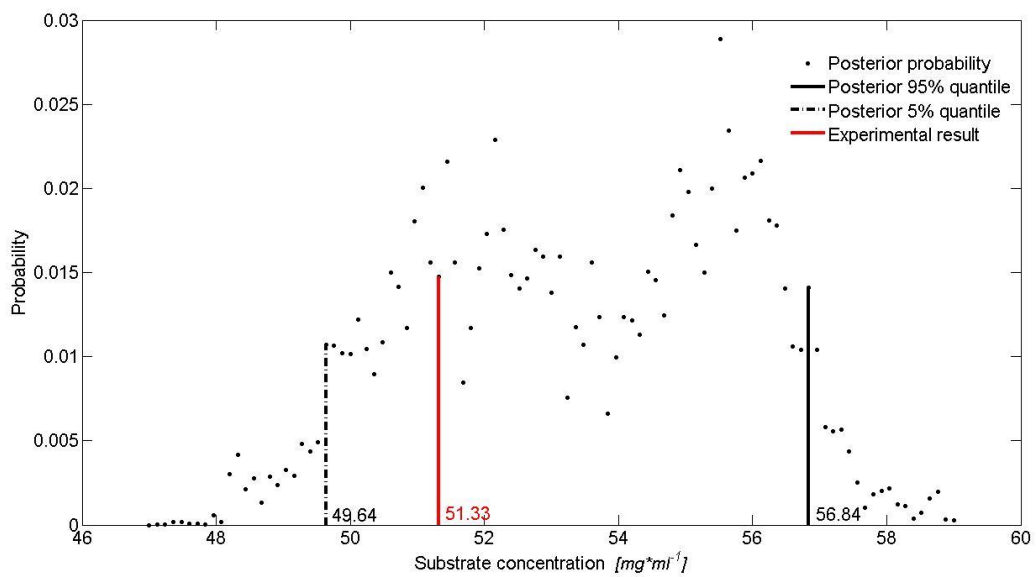
Figure E.74: Posterior probability of substrate concentration at time t=5 h Moser-Model

Annex F – Figures for 5.2.1.4

- Run 1: Application of the parameterization with the highest posterior probability as candidate vector for model discrimination

  o Concergence of Markov-Chains (exemplary)
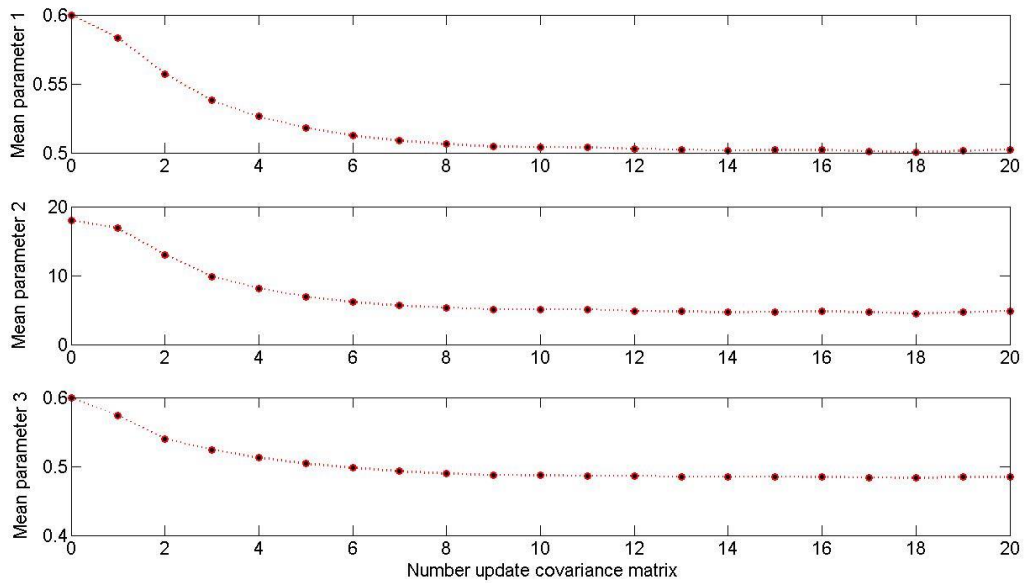
Figure F.75: Convergence mean MM-Model
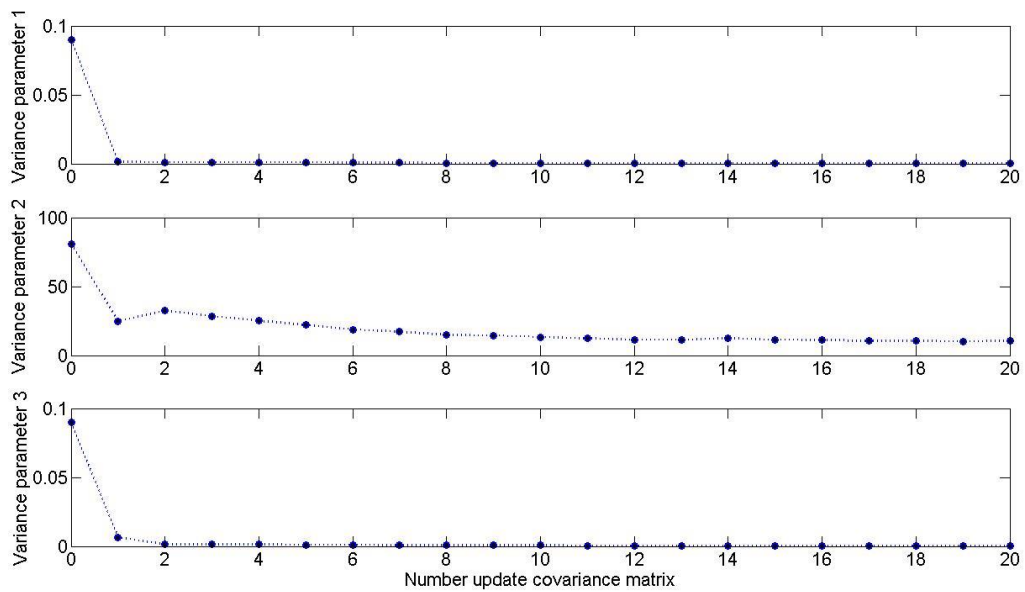
Figure F.76: Convergence variance MM-Model
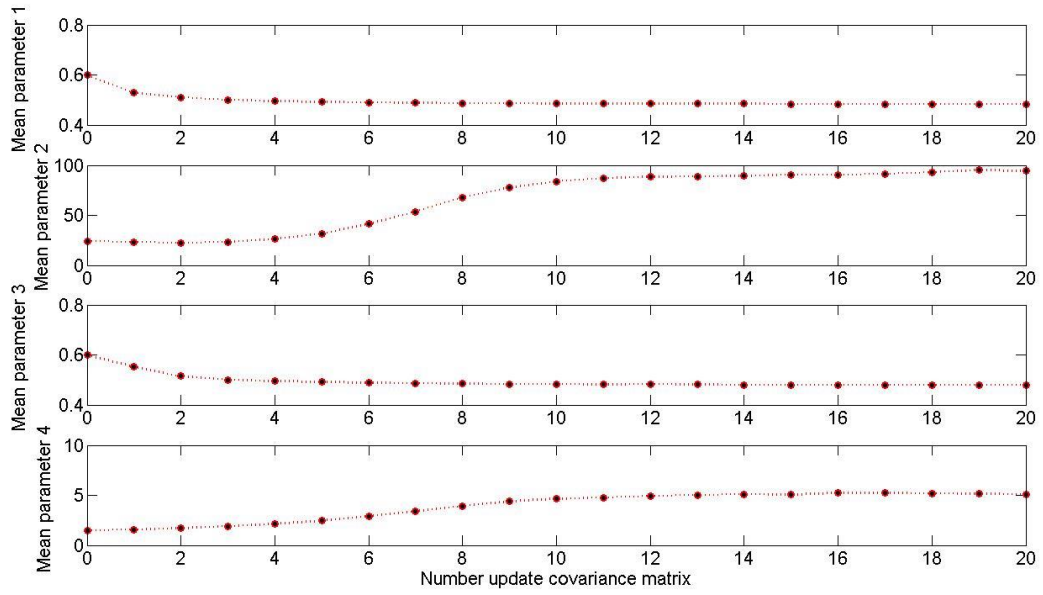
Figure F.77: Convergence mean Moser-Model
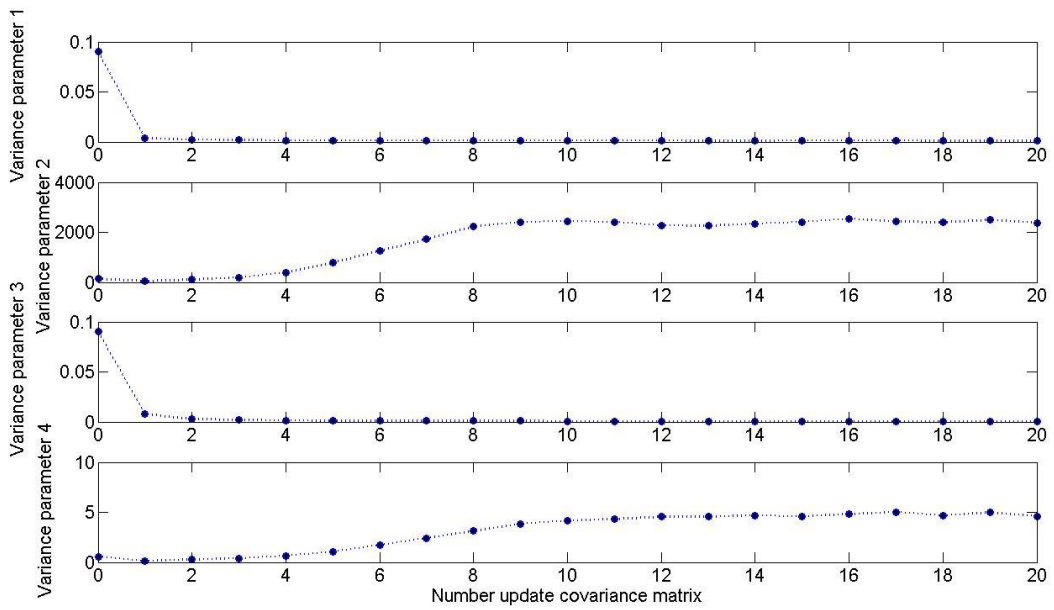
Figure F.78: Convergence variance Moser-Model

178

o   Comparison of Bayesian- and RLS-Model-Fit
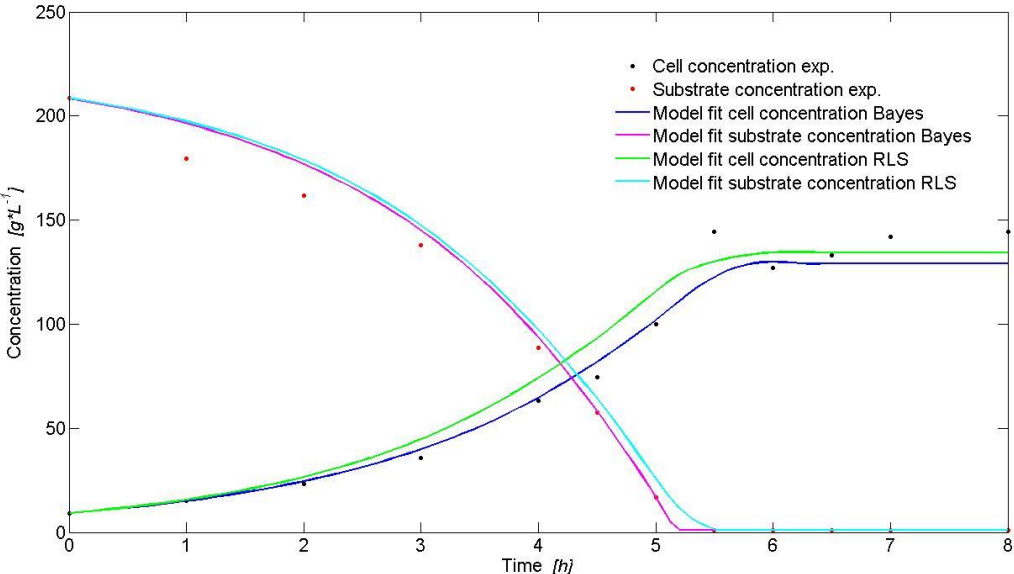
Figure F.79: Bayesian- and RLS-Model-Fit of MM-Model
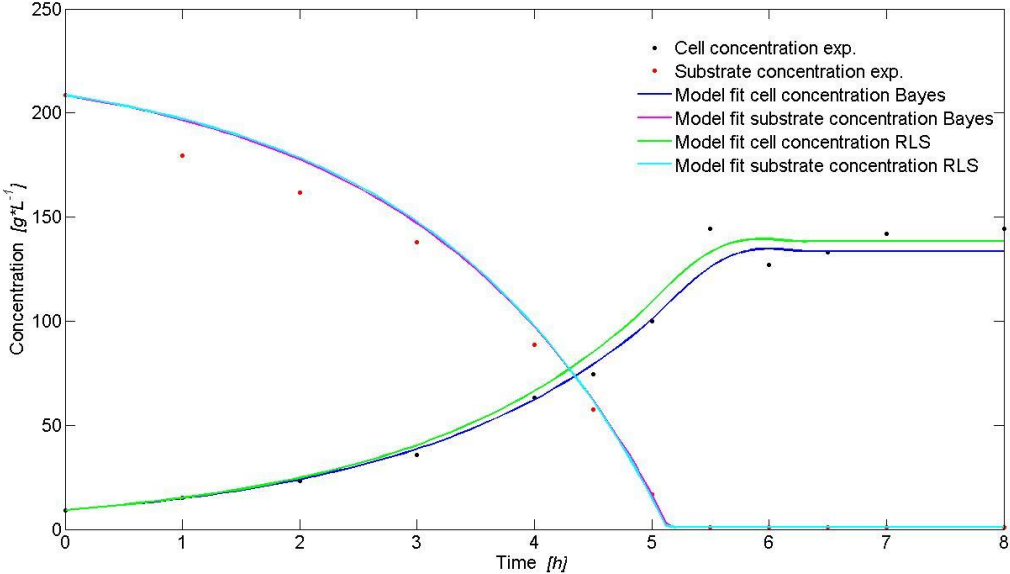
Figure F.80: Bayesian- and RLS-Model-Fit of Moser-Model

179

o Prior and posterior parameter distributions (exemplary)
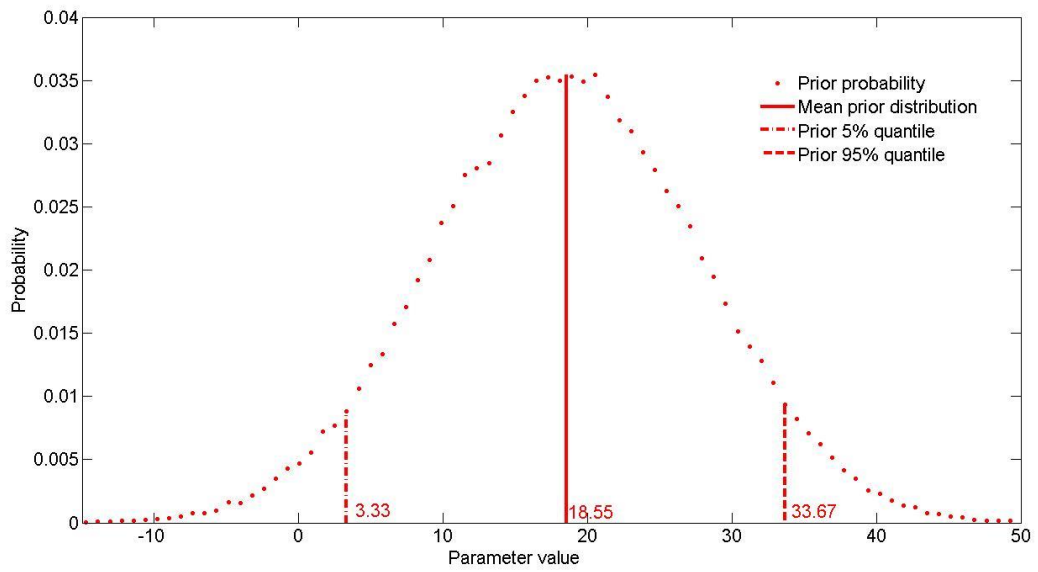
Figure F.81: Prior proabability distribution of parameter 2 MM-Model
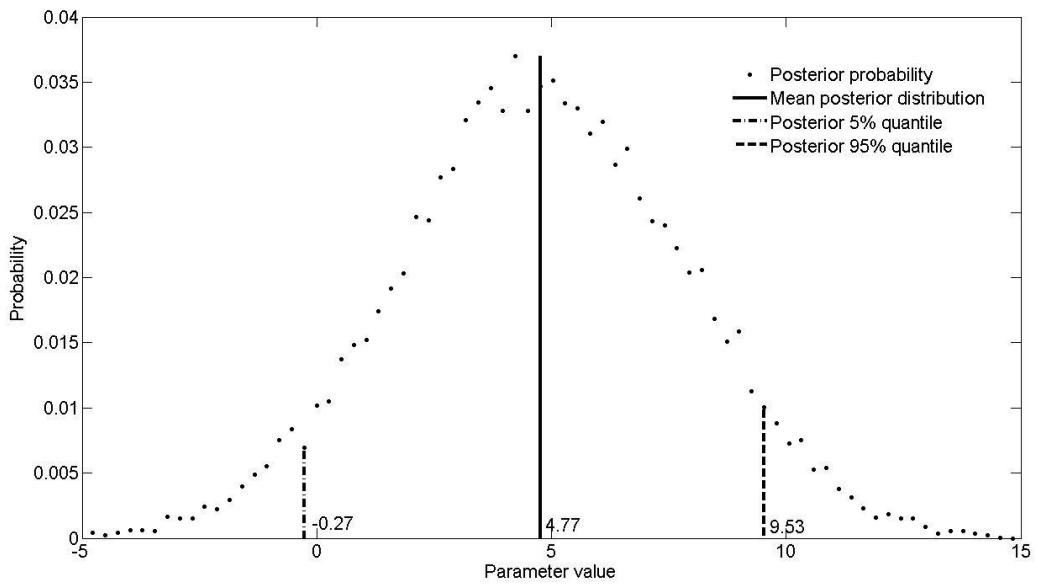
Figure F.82: Posterior proabability distribution of parameter 2 MM-Model
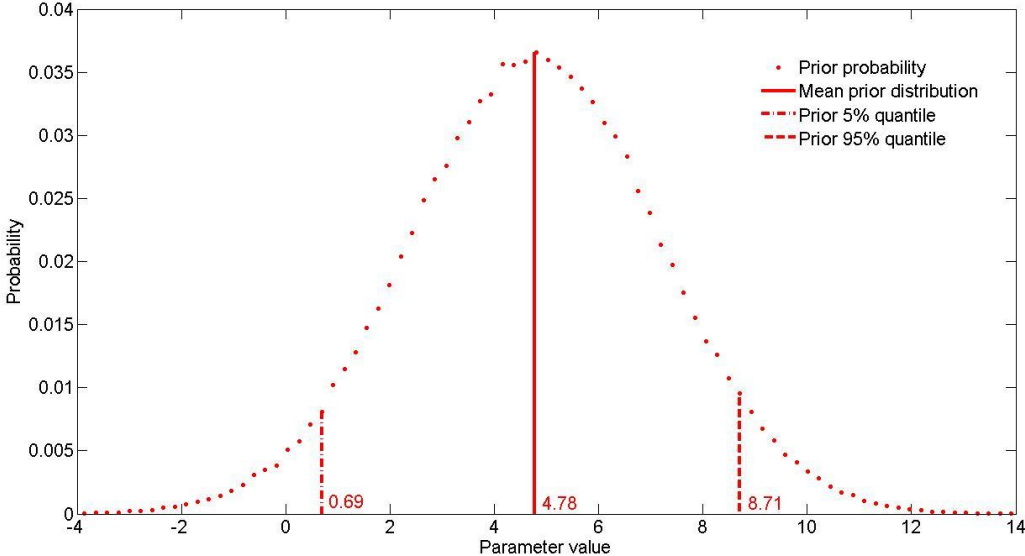
Figure F.83: Prior proability distribution of parameter 4 Moser-Model
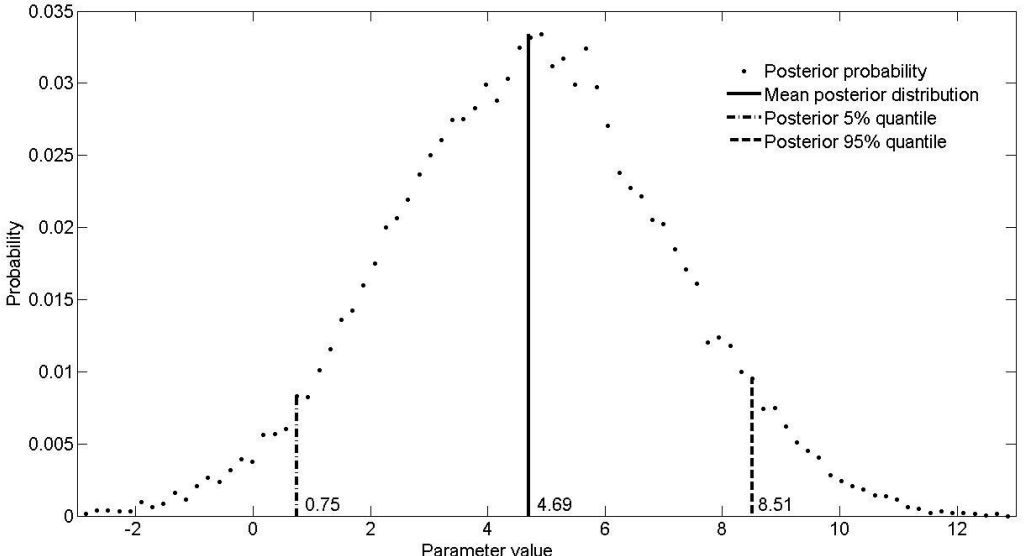
Figure F.84: Posterior proability distribution of parameter 4 Moser-Model
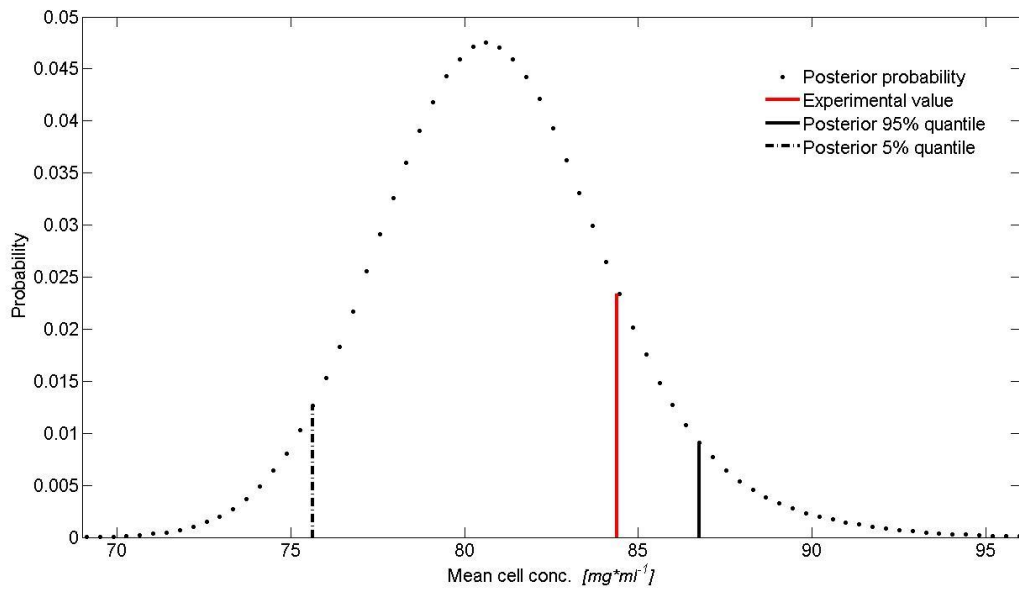
o Model validation

Figure F.85: Validation mean cell concentration MM-Model
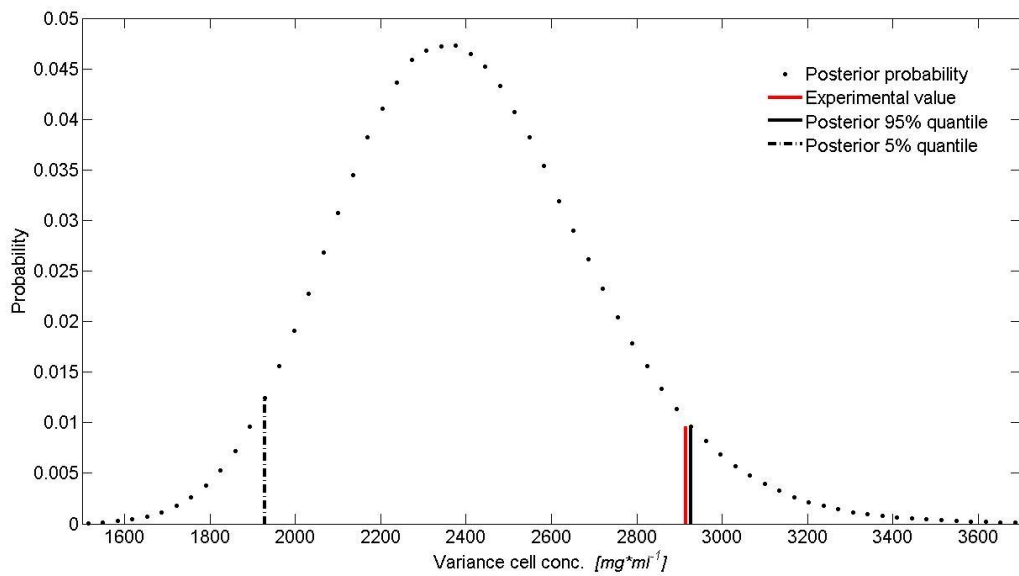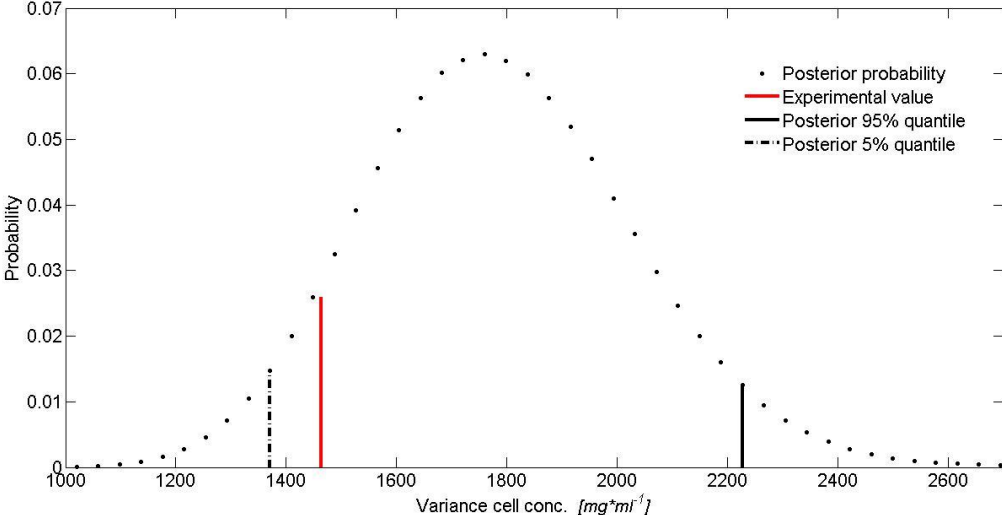

Figure F.86: Validation variance cell concentration MM-Model
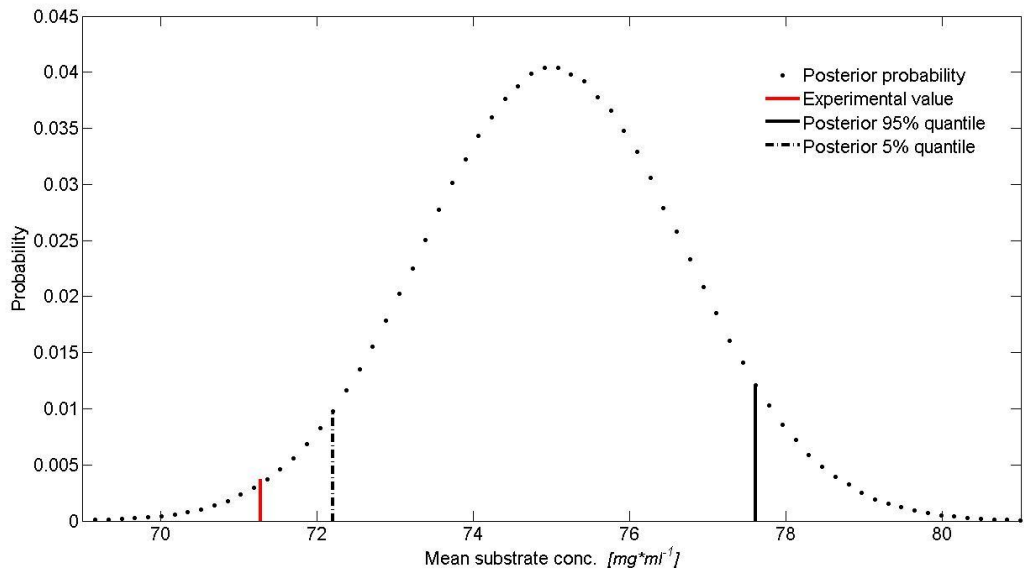
Figure F.87: Validation mean substrate concentration MM-Model
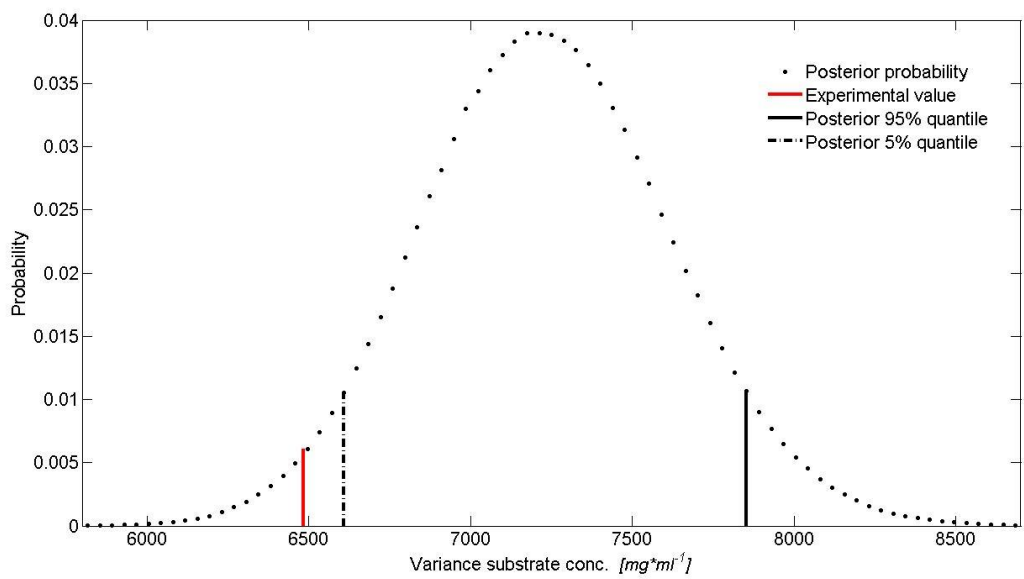
Figure F.88: Validation variance substrate concentration MM-Model
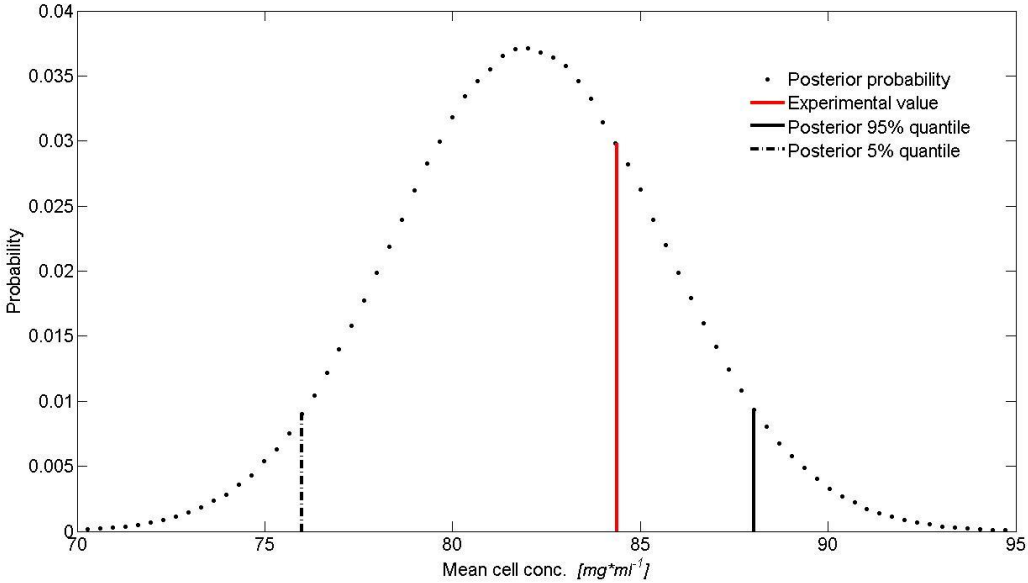
Figure F.89: Validation mean cell concentration Moser-Model
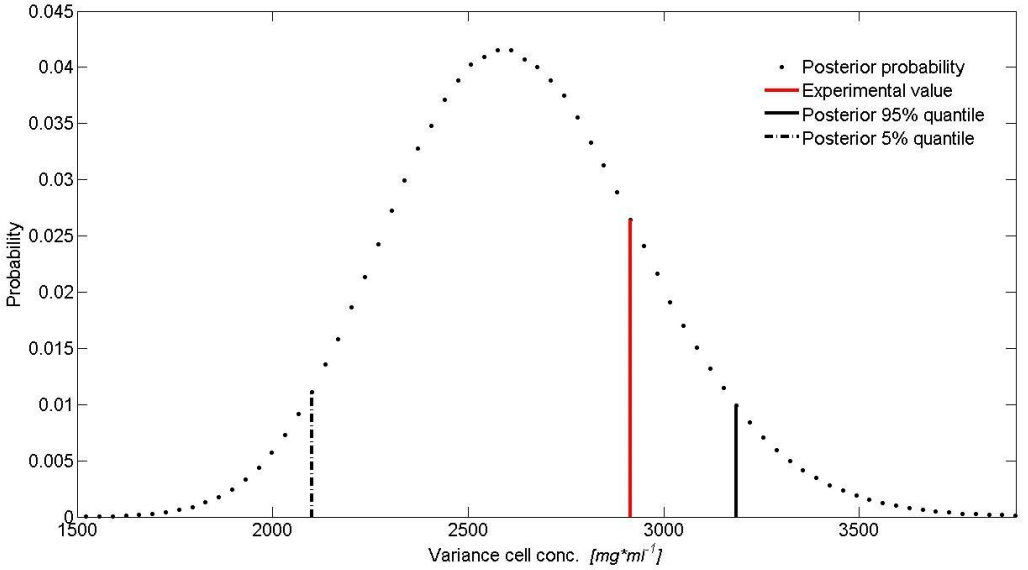
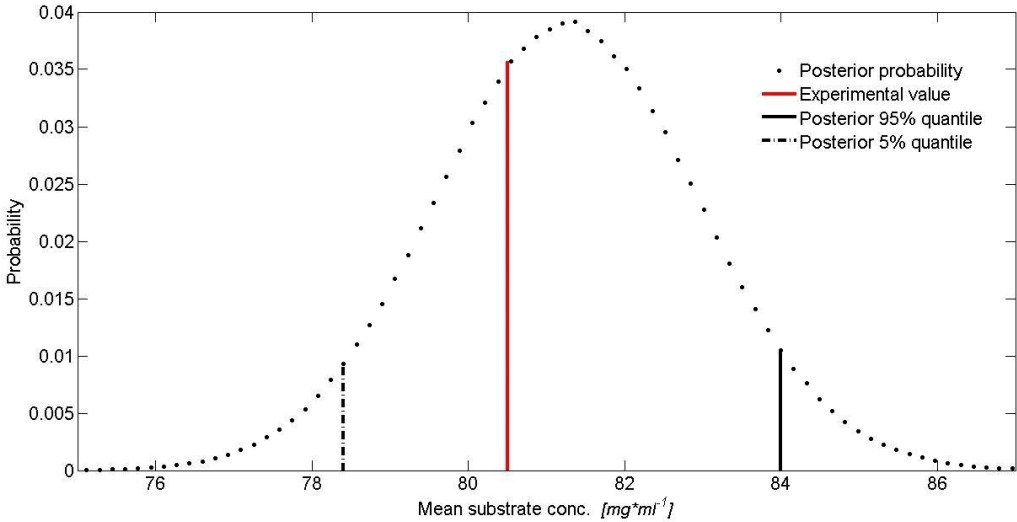Figure F.90: Validation variance cell concentration Moser-Model
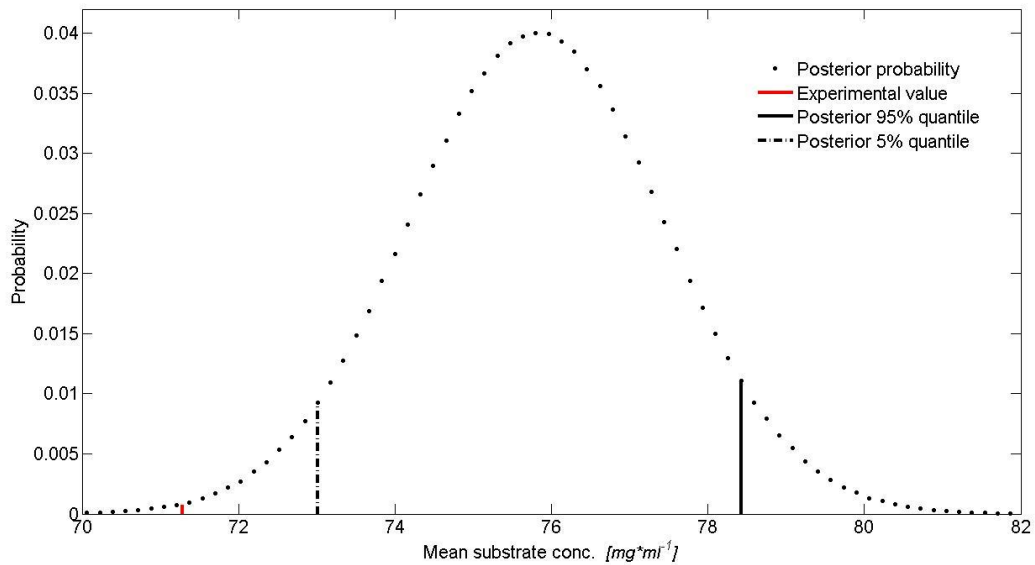
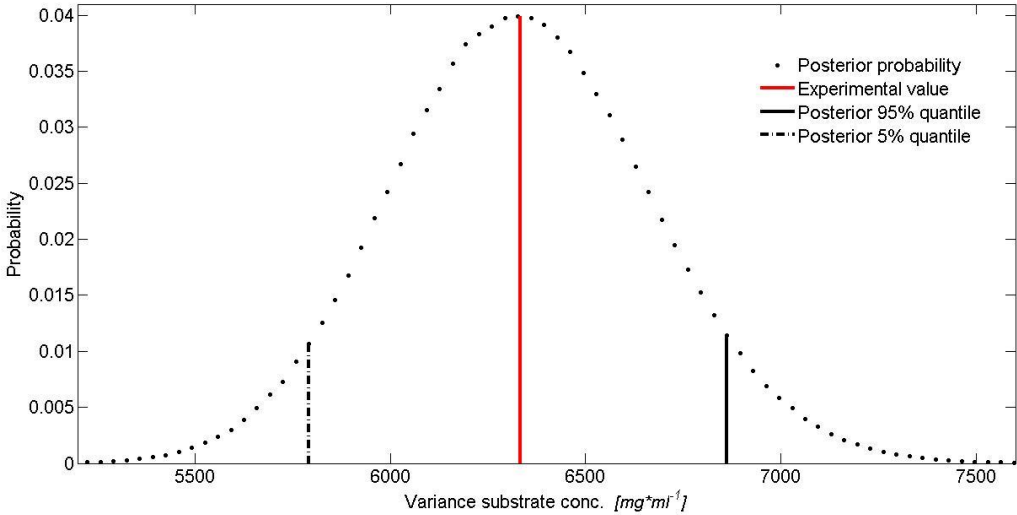Figure F.91: Validation mean substrate concentration Moser-Model
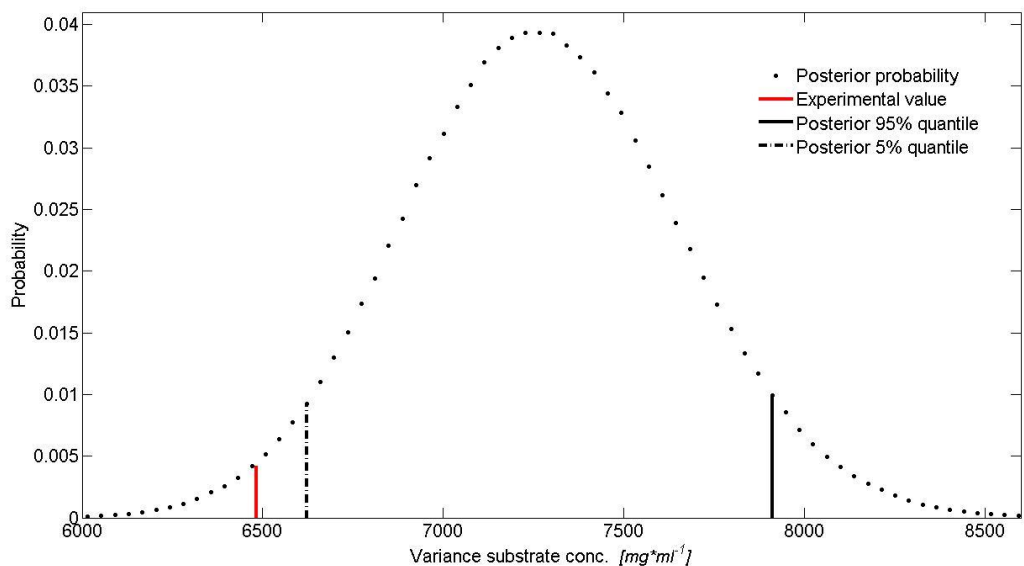
Figure F.92: Validation variance substrate concentration Moser-Model

- o Posterior probability distribution of experimental observations (exemplary)

  Not necessary as neither model is valid. Thus experimental values are expected to be not in support of distributions.

- Run 2: Application of the posterior mean as candidate vector for model discrimination

The figures exposed for Run 1 are also valid here as these results are not dependent on the choice of the candidate vector. The results only differ for the relative model probability (compare Table 5.23).

185

Annex G – Figures for 5.2.1.5 and 5.2.1.6

- Simulation 'Exp. data (Run 1)': The figures / results exposed here do not depend on the choice of the candidate vector for model discrimination. They are also identical for simulation 'Exp. data (Run 2)' which is a repetition of the simulation at identical conditions. The parameter distributions are not exposed here as this has been already done exemplarily for the other simulation experiments. The characteristic measures for comparison of the prior and posterior parameter distribution can be taken from Table 5.25. A graphical exposition does not provide any additional information here. Also the distribution of experimental observations is not relevant for the analysis of the real experimental data as the basic objective here is the model fit and discrimination, as well as its dependency on the prior variance. The analysis and discussion of experimental observation distributions has been achieved sufficiently for some examples in the preceeding simulation experiments. There is no additional use to achieve that here again.

  o Concergence of Markov-Chains (exemplary)
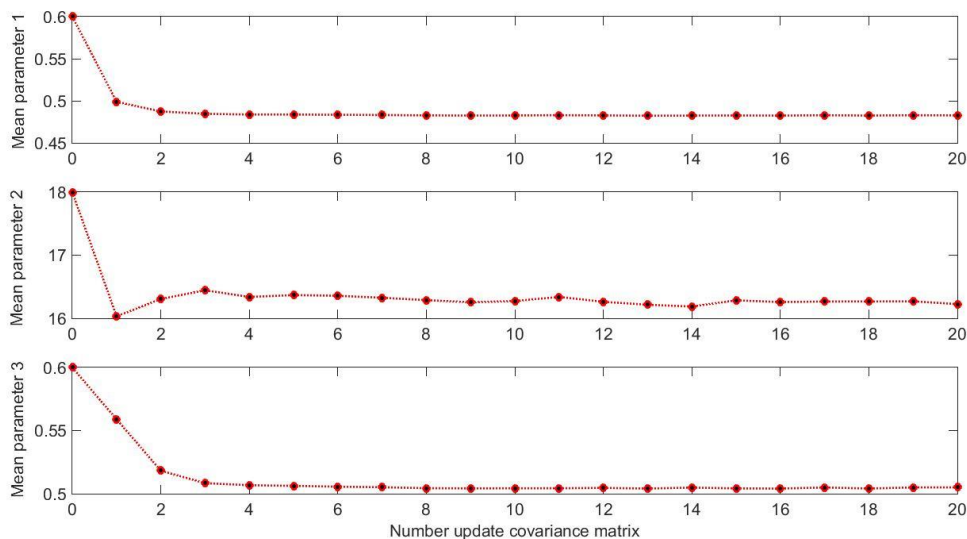
Figure G.93: Convergence mean MM-Model
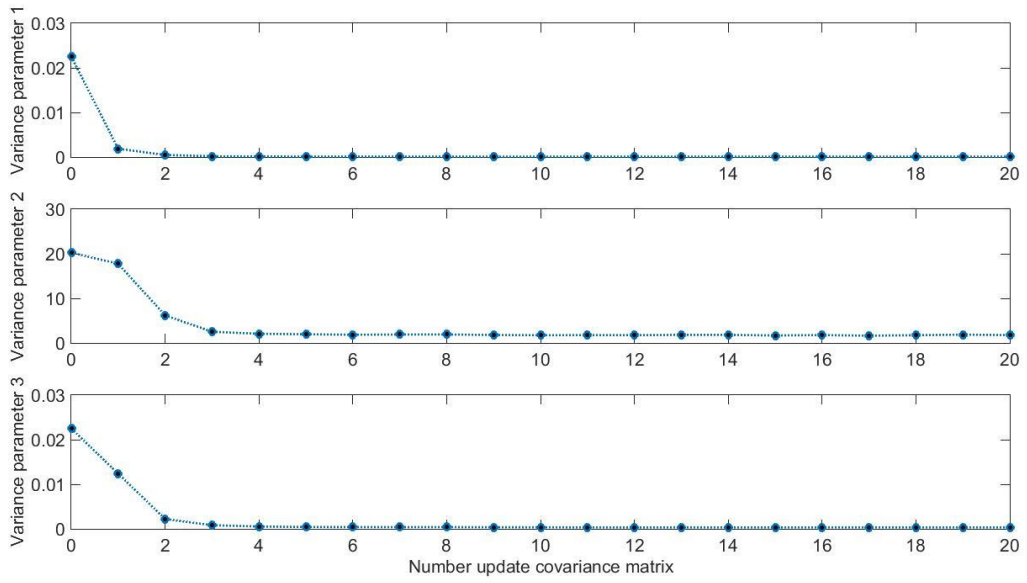
Figure G.94: Convergence variance MM-Model
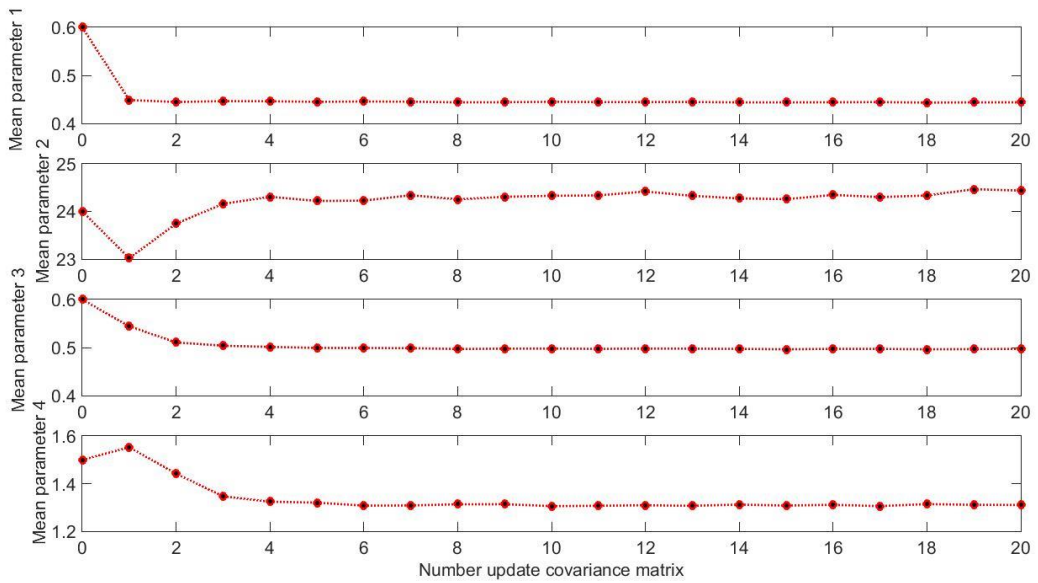
Figure G.95: Convergence mean Moser –Model
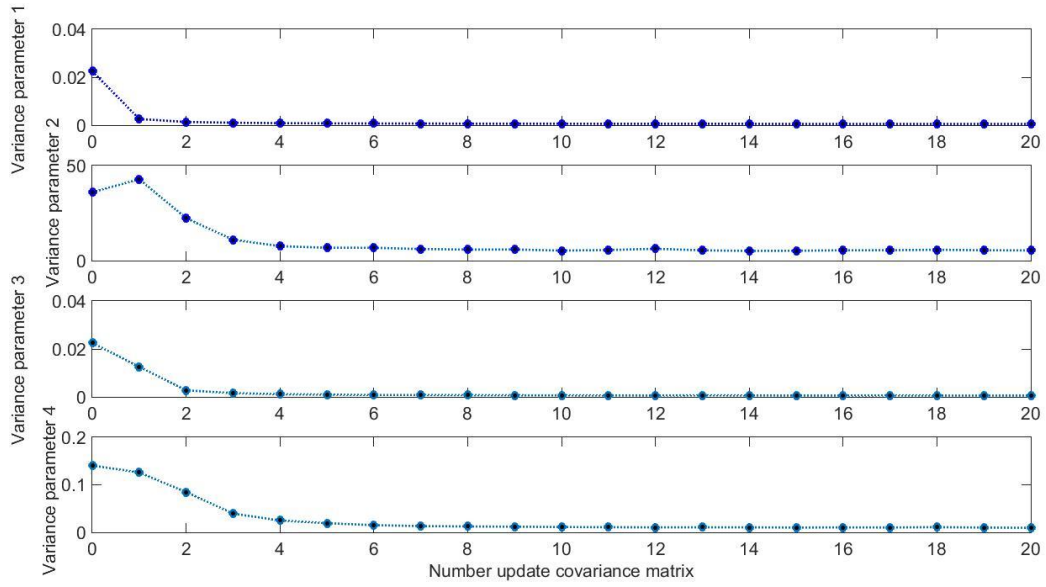
Figure G.96: Convergence variance Moser –Model

o   Comparison of Bayesian- and RLS-Model-Fit
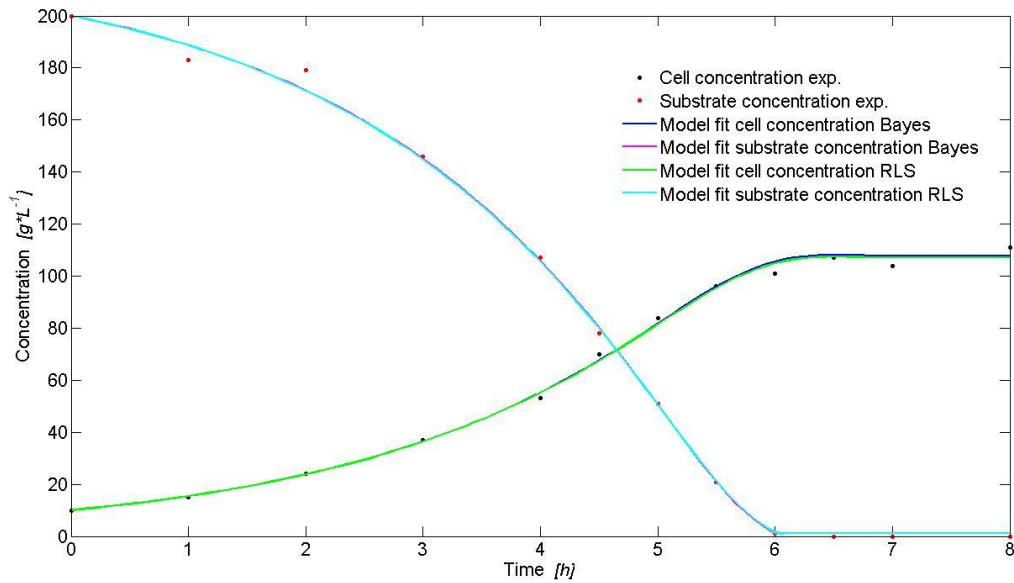
Figure G.97: Bayesian- and RLS-Model-Fit of MM-Model
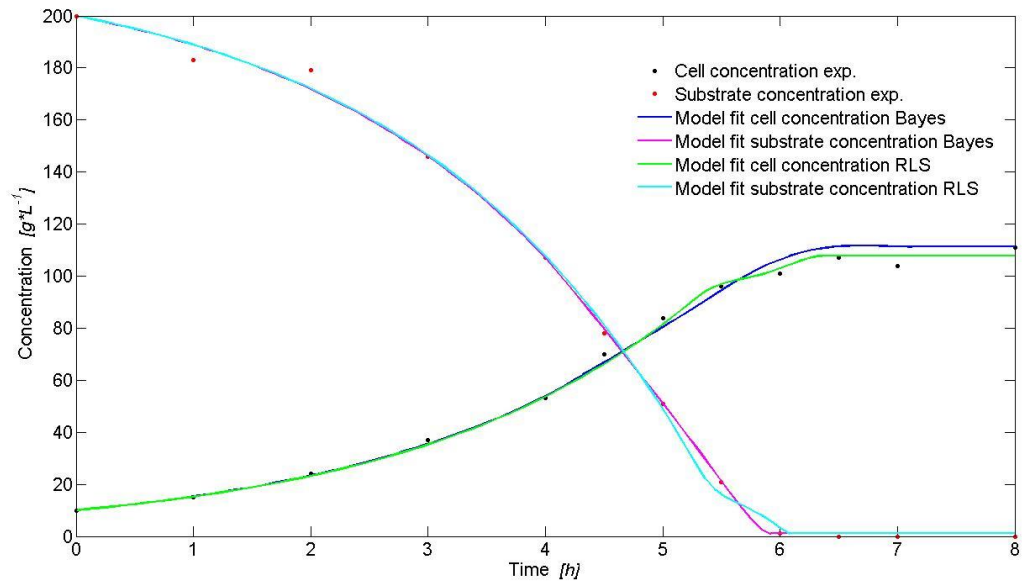
Figure G.98: Bayesian- and RLS-Model-Fit of Moser-Model
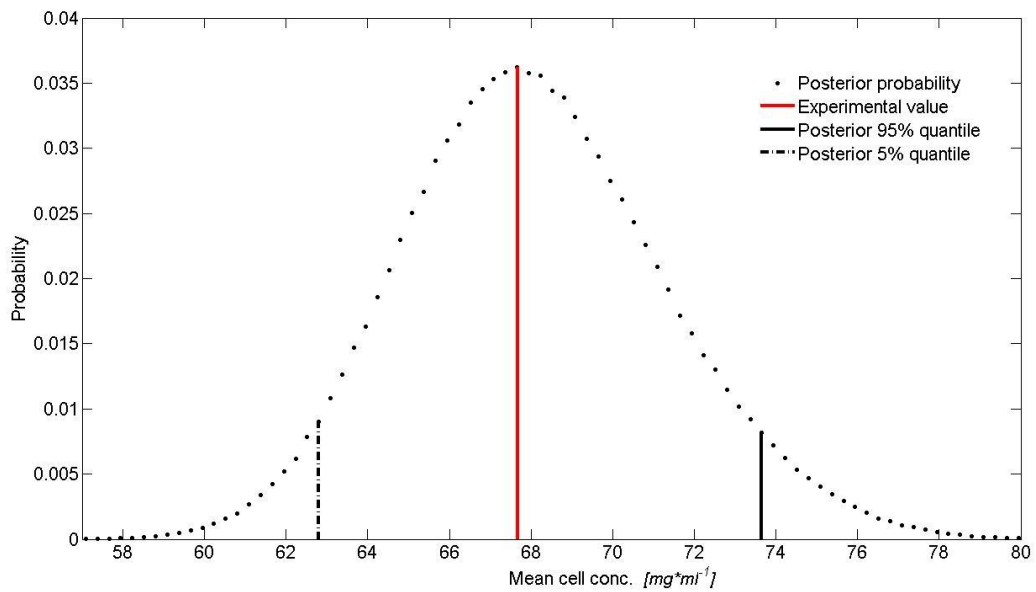
o   Model validation

Figure G.99: Validation mean cell concentration MM-Model
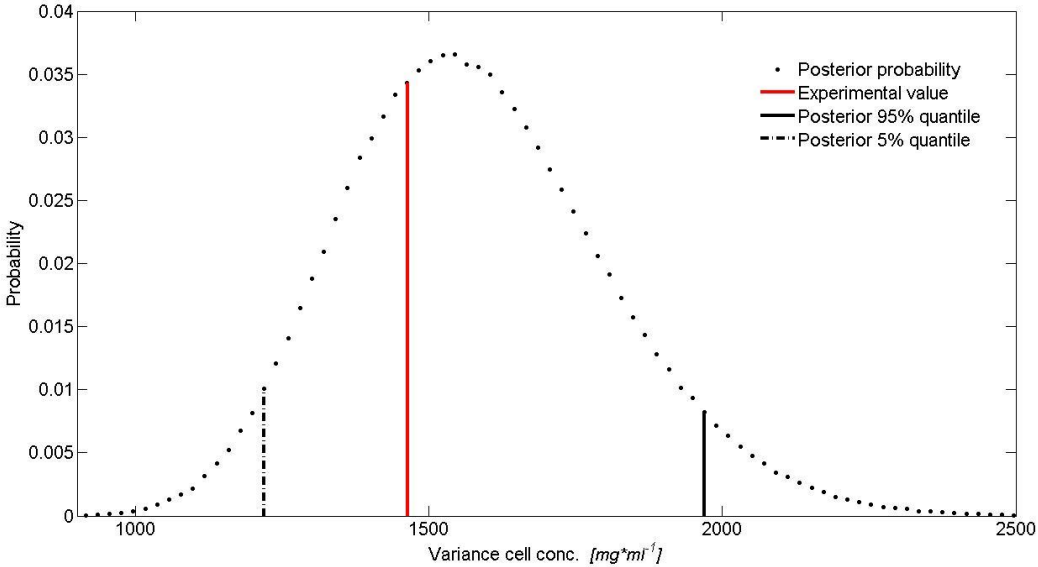
Figure G.100: Validation variance cell concentration MM-Model
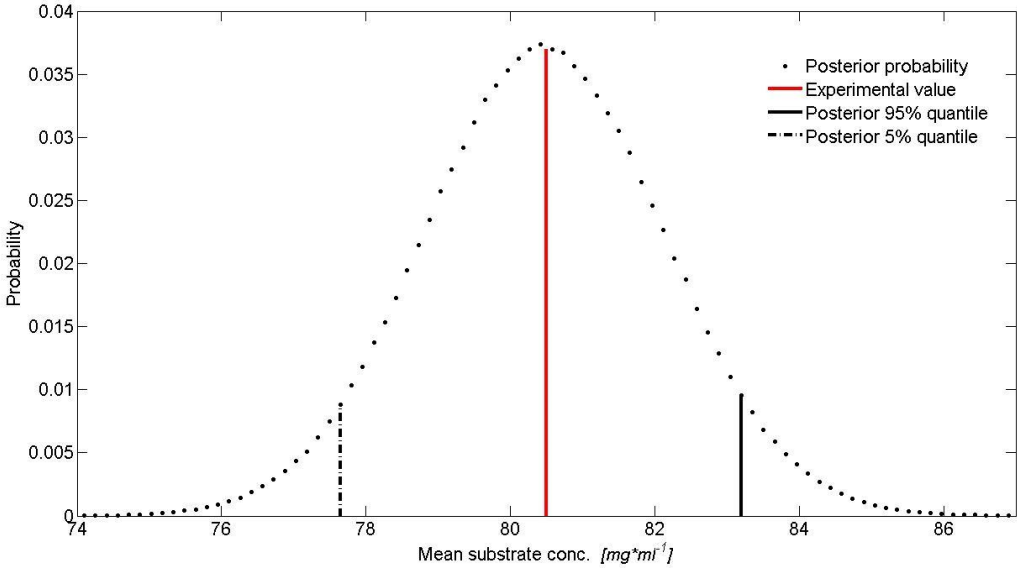
Figure G.101: Validation mean substrate concentration MM-Model
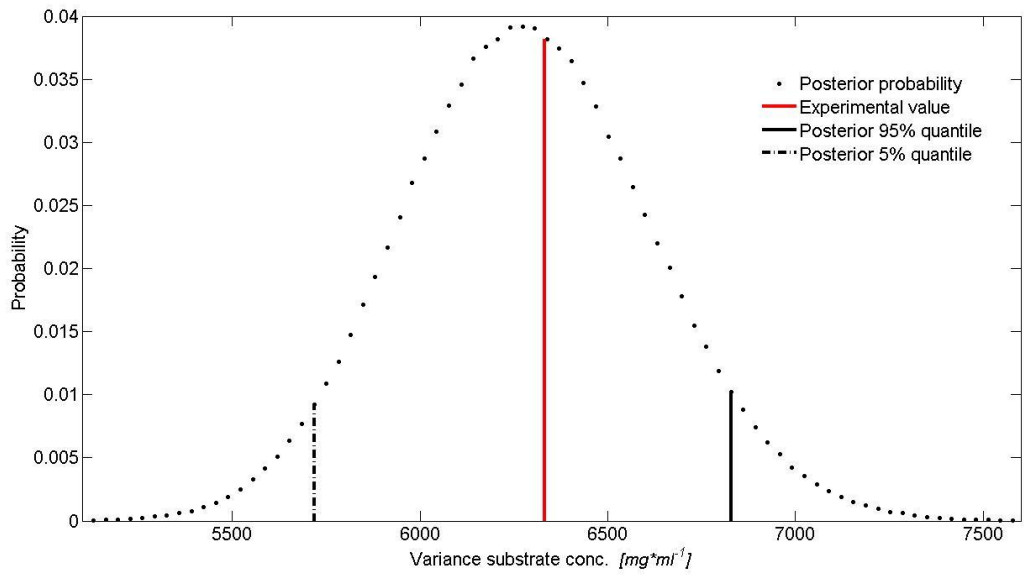
Figure G.102: Validation variance substrate concentration MM-Model
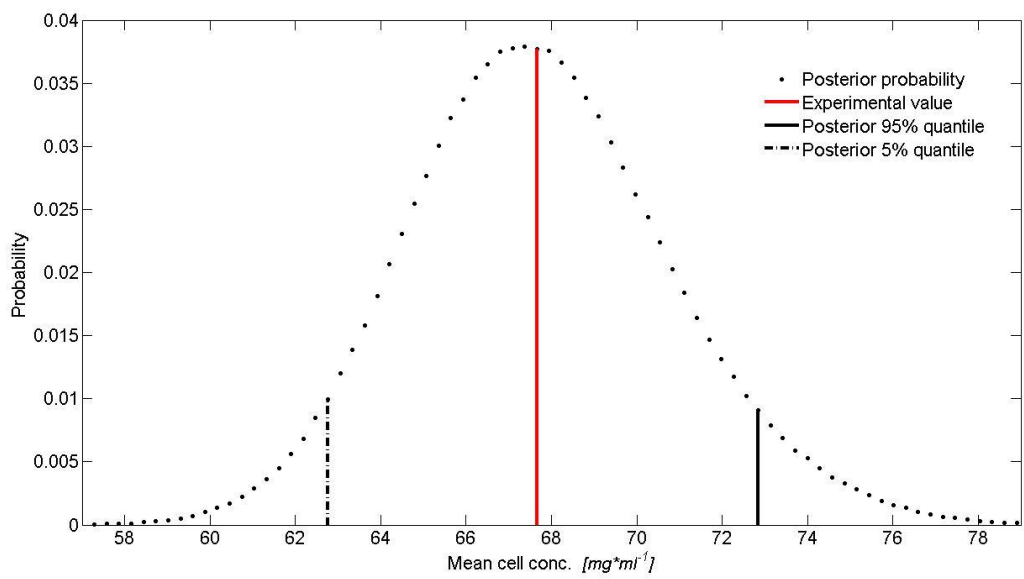
Figure G.103: Validation mean cell concentration Moser-Model

Figure G.104: Validation variance cell concentration Moser-Model

Figure G.105: Validation mean substrate concentration Moser-Model

Figure G.106: Validation variance substrate concentration Moser-Model

- Simulation 'Exp. data (Run 3)': The figures / results exposed here do not depent on the choice of the candidate vector for model discrimination. The parameter distributions are not exposed here as this has been already done exemplarily for the other simulation experiments. The characteristic measures for comparison of the prior and posterior parameter distribution can be taken from Table 5.25. A graphical exposition does not provide any additional information here. Also the distribution of experimental observations is not relevant for the analysis of the real experimental data as the basic objective is the model fit, discrimination and validation here as well as its dependency on the prior variance. The analysis and discussion of experimental observation distributions has been achieved sufficiently for some examples in the preceeding simulation experiments. There is no additional use to achieve that here again.

o   Concergence of Markov-Chains (exemplary)
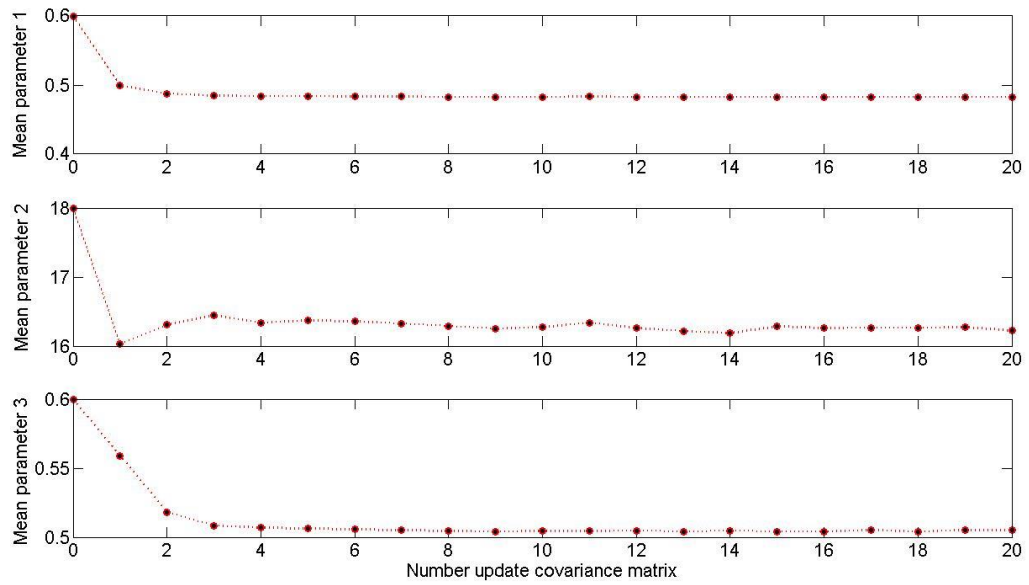
Figure G.107: Convergence mean Michaelis-Menten–Model

Figure G.108: Convergence variance Michaelis-Menten–Model

Figure G.109: Convergence mean Moser–Model

Figure G.110: Convergence variance Moser–Model

o   Comparison Bayesian- and RLS-Model-Fit
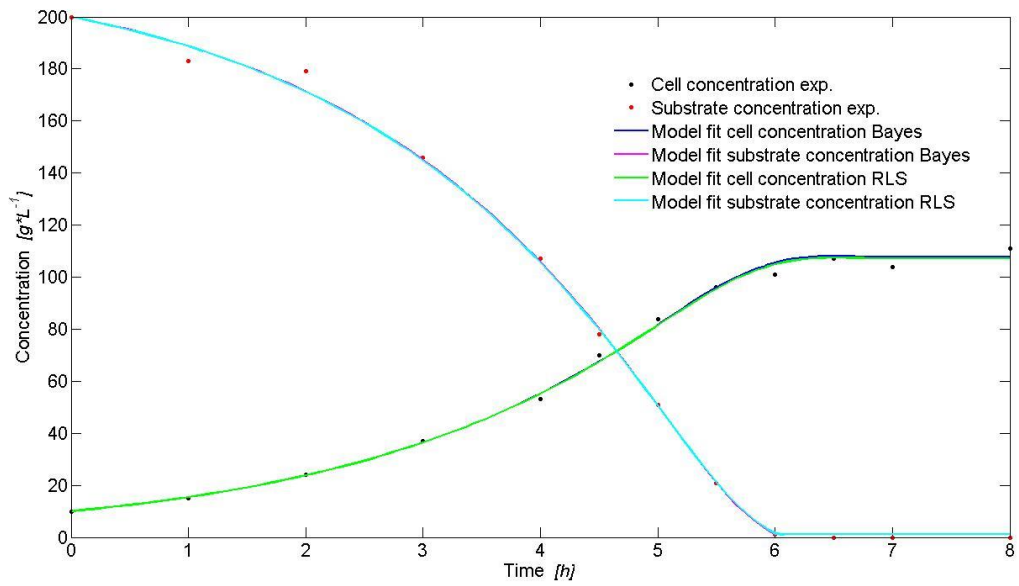
Figure G.111: Bayesian- and RLS-Model-Fit of MM-Model
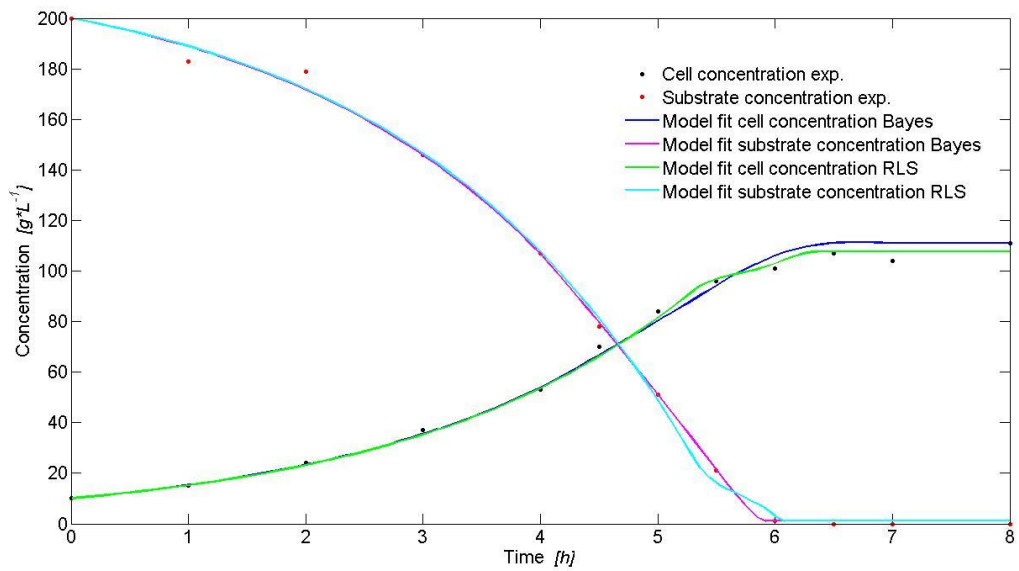
Figure G.112: Bayesian- and RLS-Model-Fit of Moser-Model

o   Model validation

Figure G.113: Validation mean cell concentration MM-Model

Figure G.114: Validation variance cell concentration MM-Model
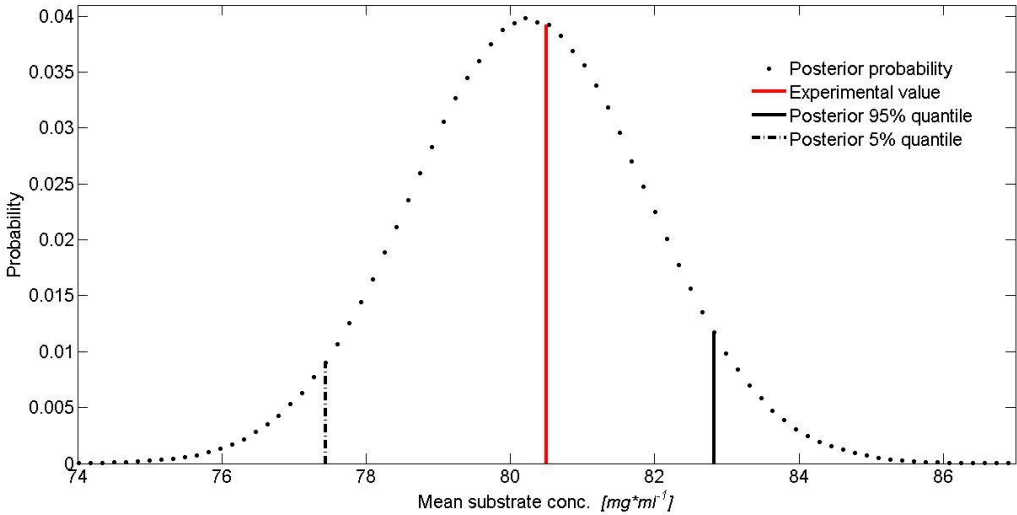
Figure G.115: Validation mean substrate concentration MM-Model
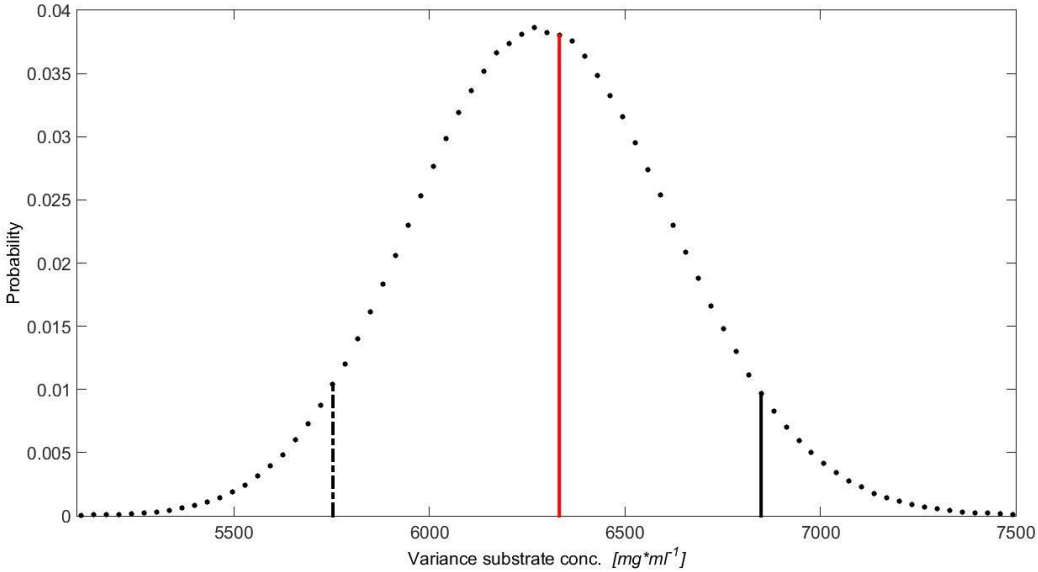
Figure G.116: Validation variance substrate concentration MM-Model
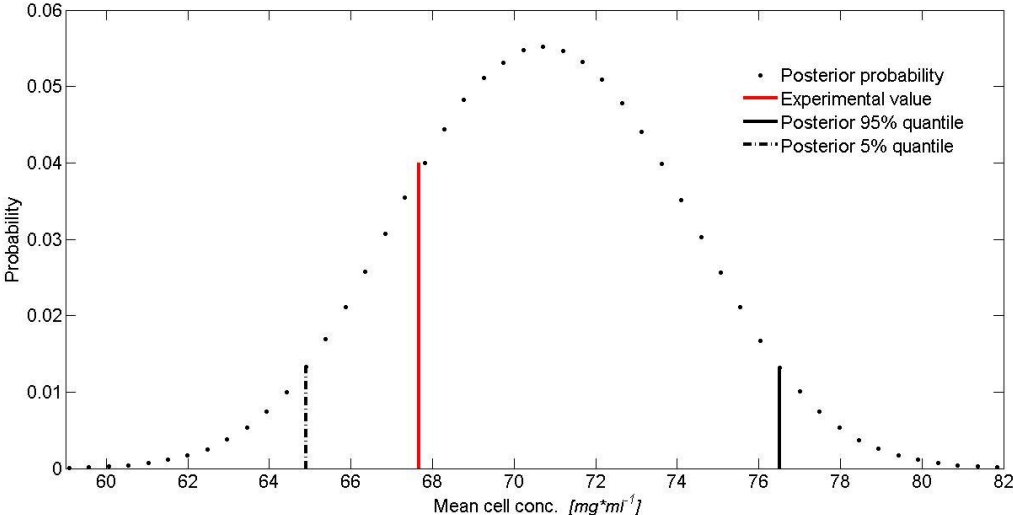
Figure G.117: Validation mean cell concentration Moser-Model

Figure G.118: Validation variance cell concentration Moser-Model

Figure G.119: Validation mean substrate concentration Moser-Model

Figure G.120: Validation variance substrate concentration Moser-Model

Annex H – Figures for 5.2.2.1

  o Concergence of Markov-Chains (exemplary)
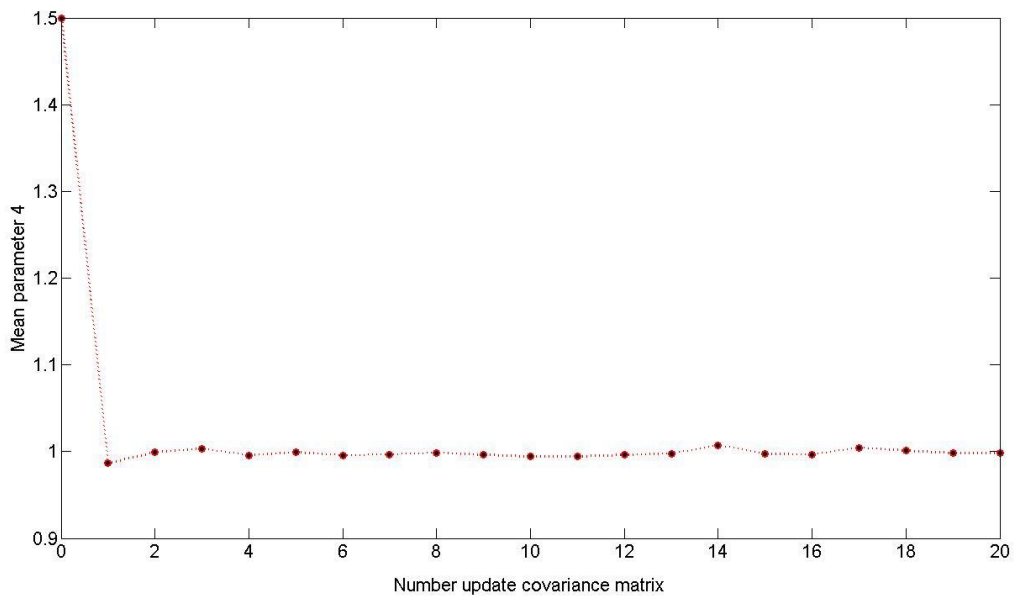
Figure H.121: Convergence mean parameter 4 model 1st grade rancimat value
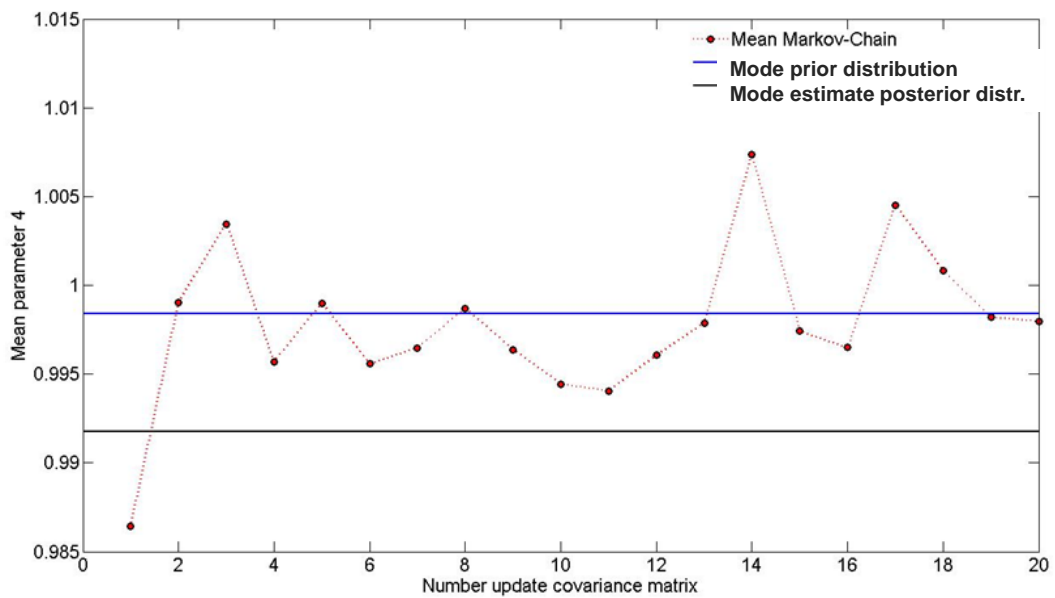
Figure H.122: Convergence mean parameter 4 model 1st grade rancimat value (zoom)

201
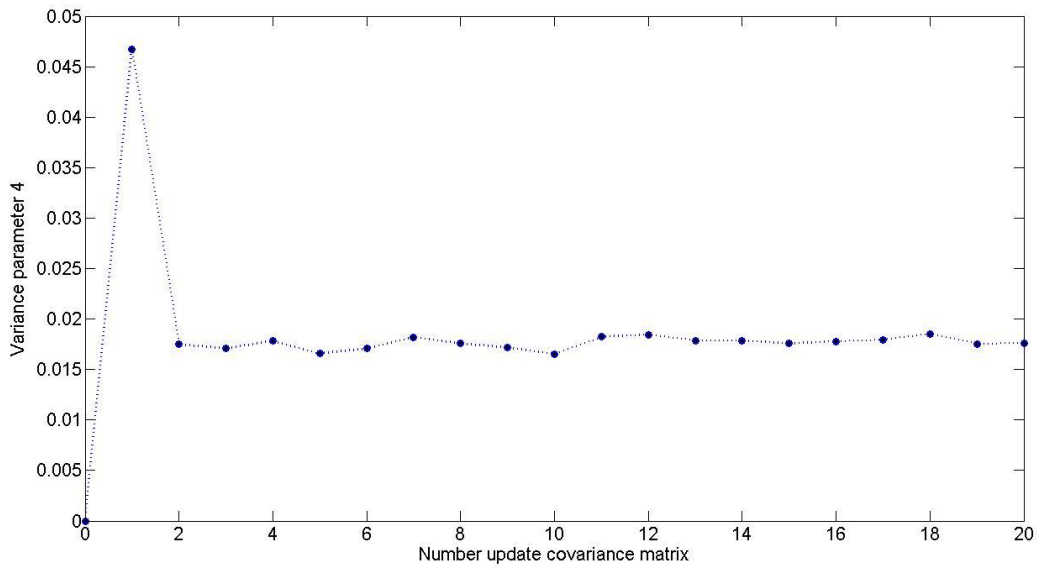
Figure H.123: Convergence variance parameter 4 model 1st grade rancimat value
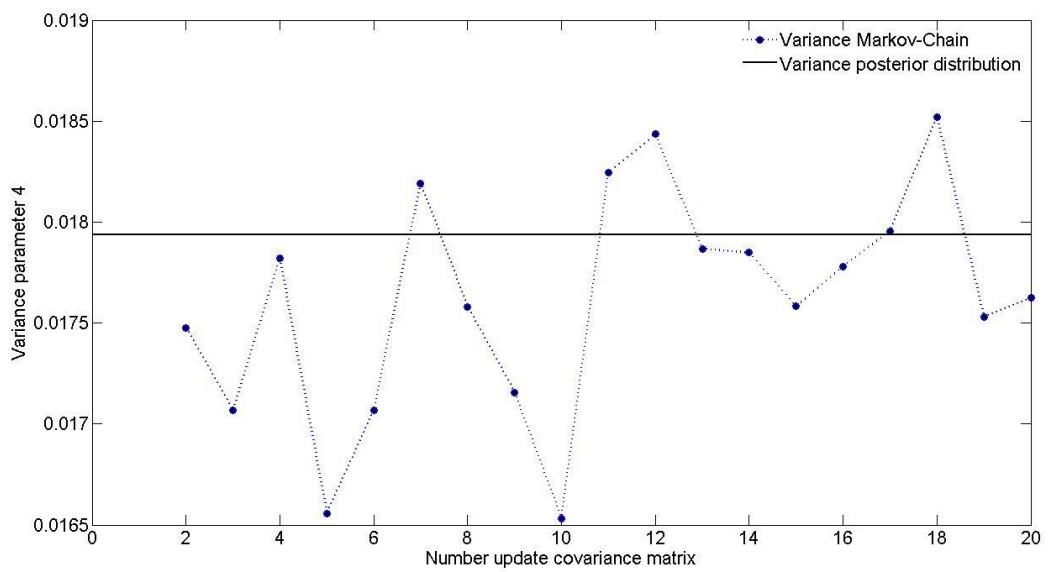
Figure H.124: Convergence variance parameter 4 model 1st grade rancimat value (zoom)

o Bayesian model fit
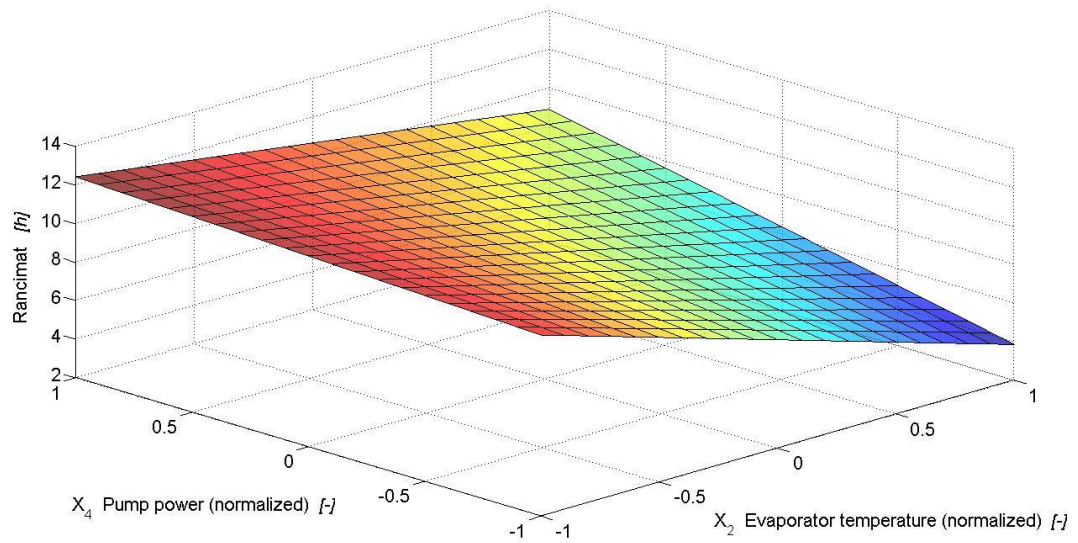
Figure H.125: Model 1st grade rancimat value (Bayesian fit)

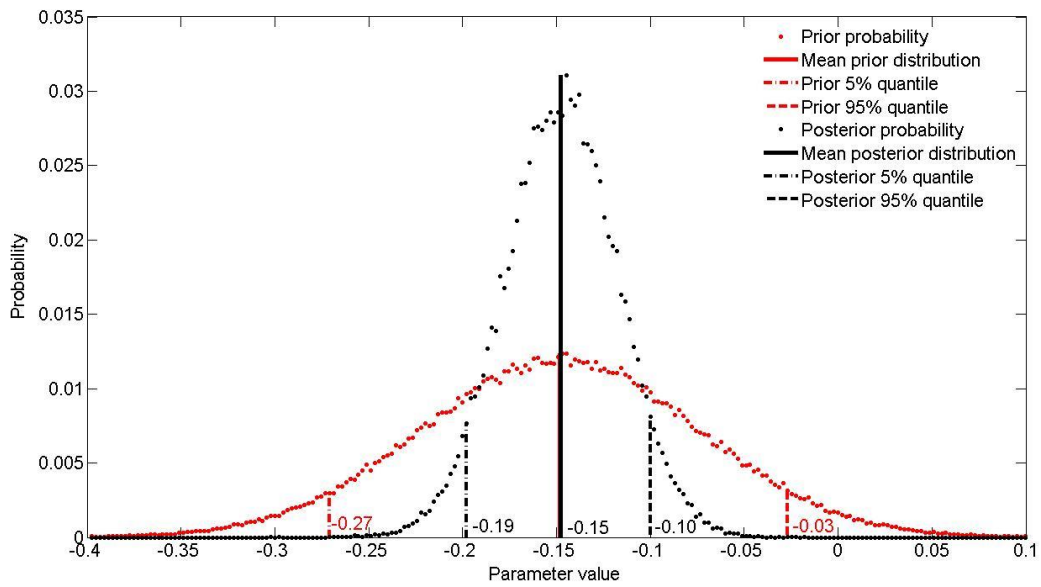o Prior and posterior parameter distributions (exemplary)
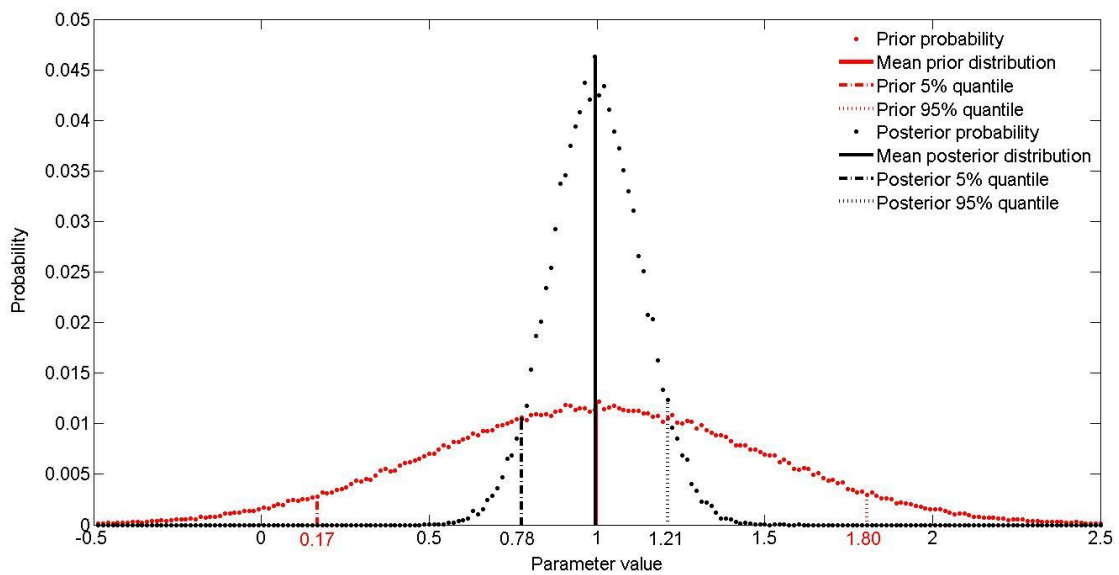
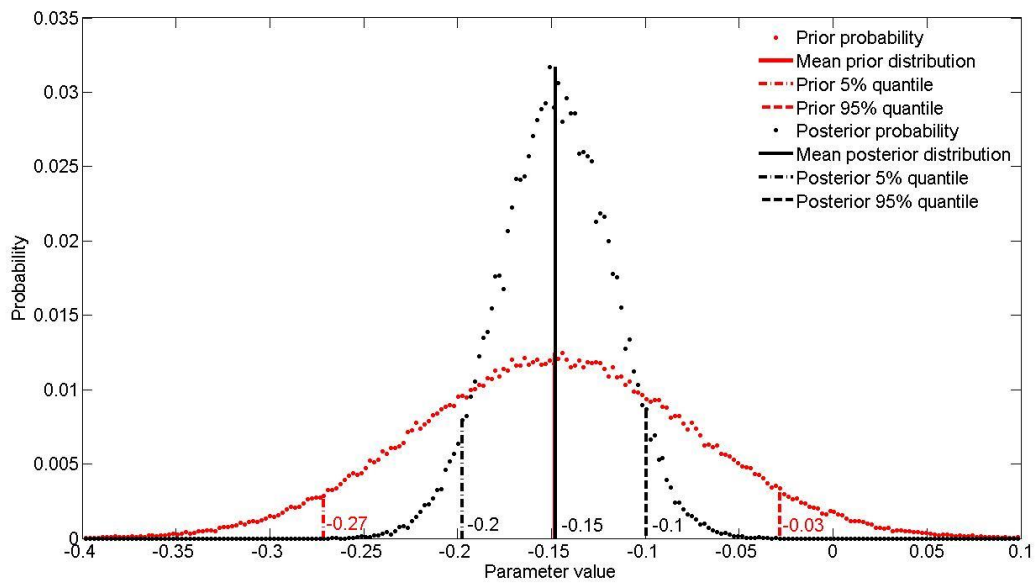Figure H.126: Prior and posterior probability distribution of parameter 4 model 1st grade rancimat value
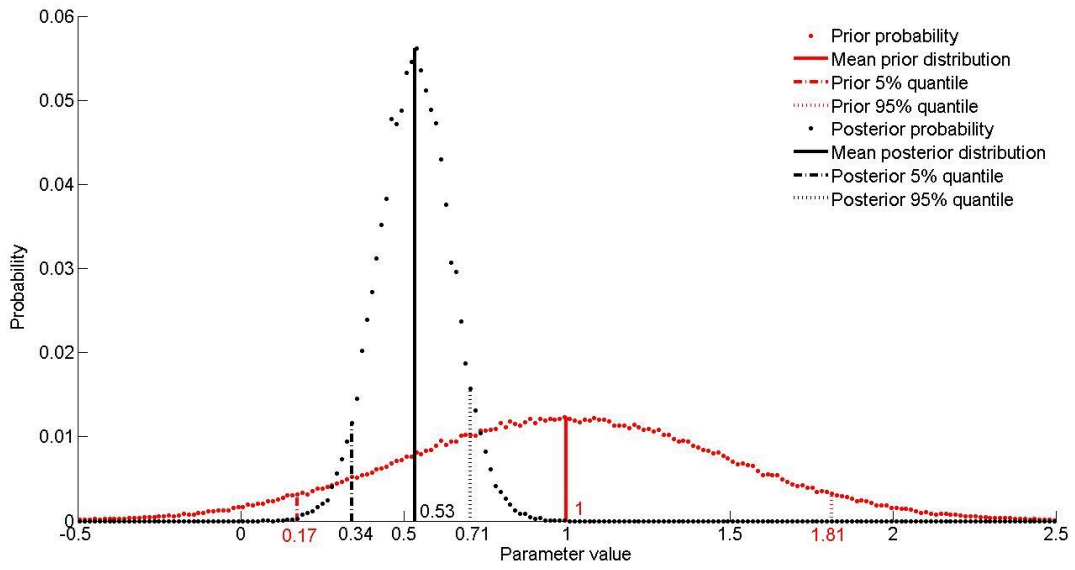
Figure H.127: Prior and posterior probability distribution parameter 9 model 2nd grade rancimat value

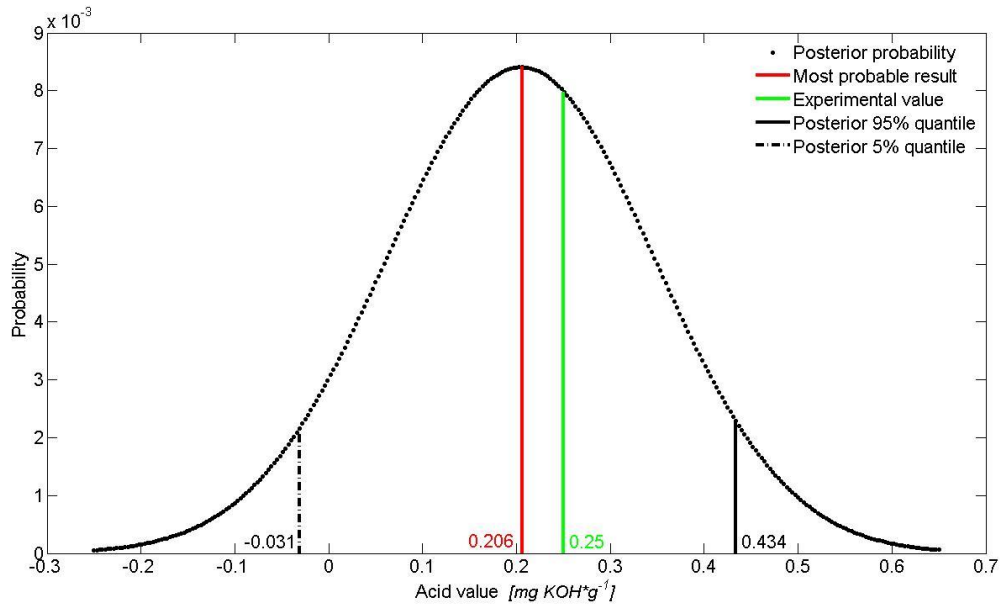o Posterior probability distribution of experimental observations (exemplary)
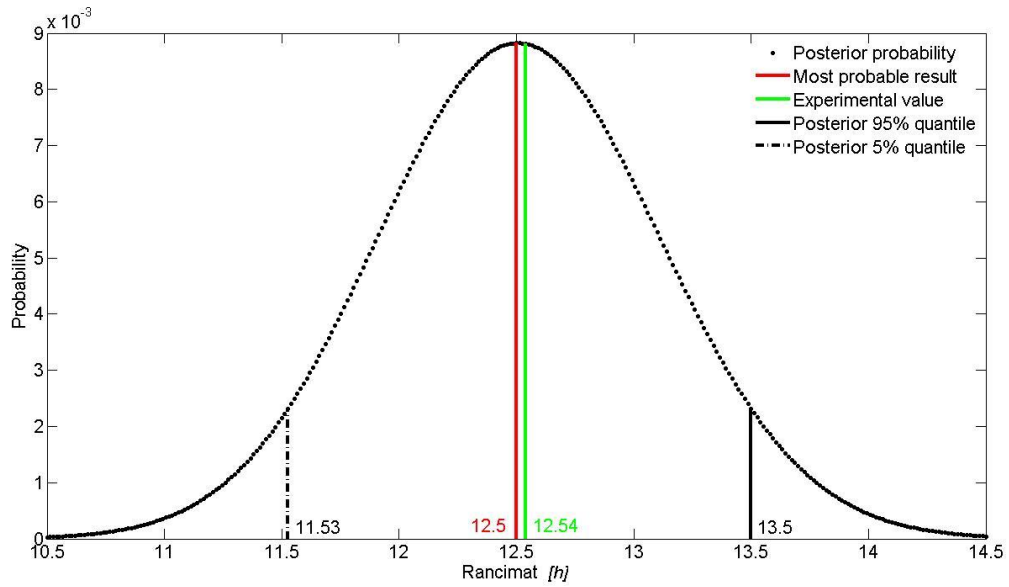
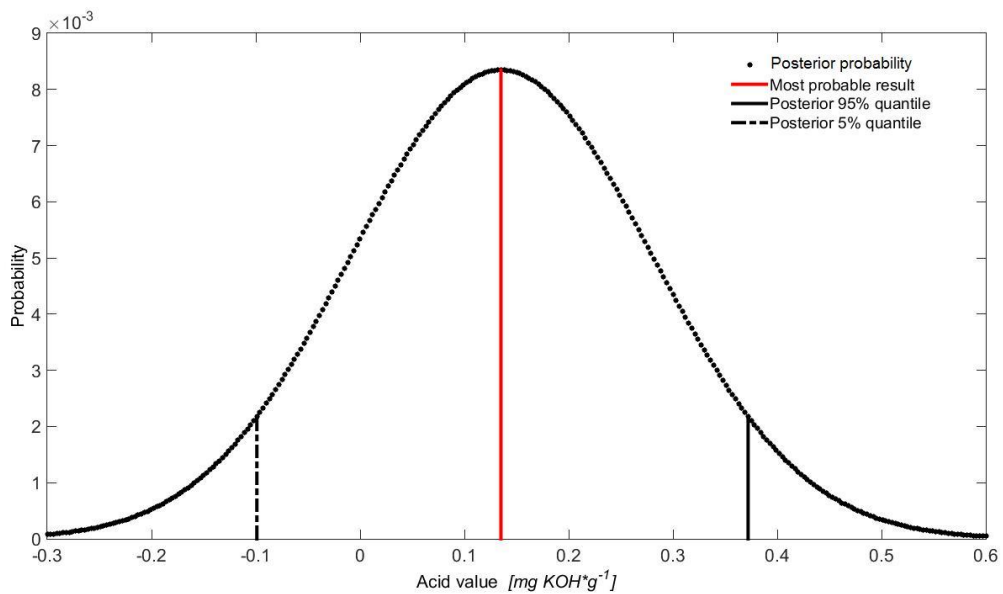Figure H.128: Posterior probability of exp. observations at opt. setting model 1st grade rancimat value
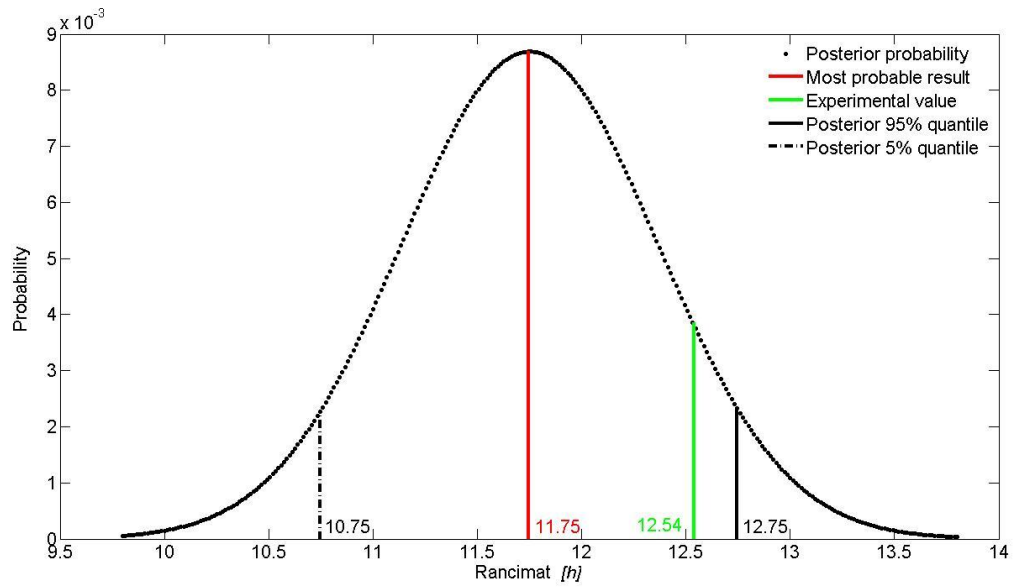
Figure H.129: Posterior probability of exp. observations at opt. setting model 2nd grade rancimat value

o   Model validation
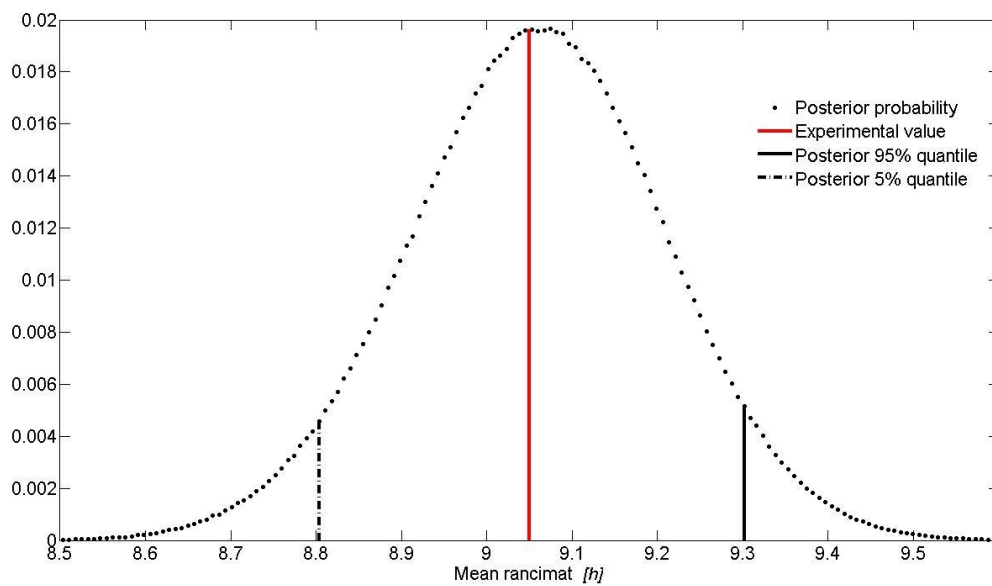
Figure H.130: Validation mean rancimat value model 1st grade
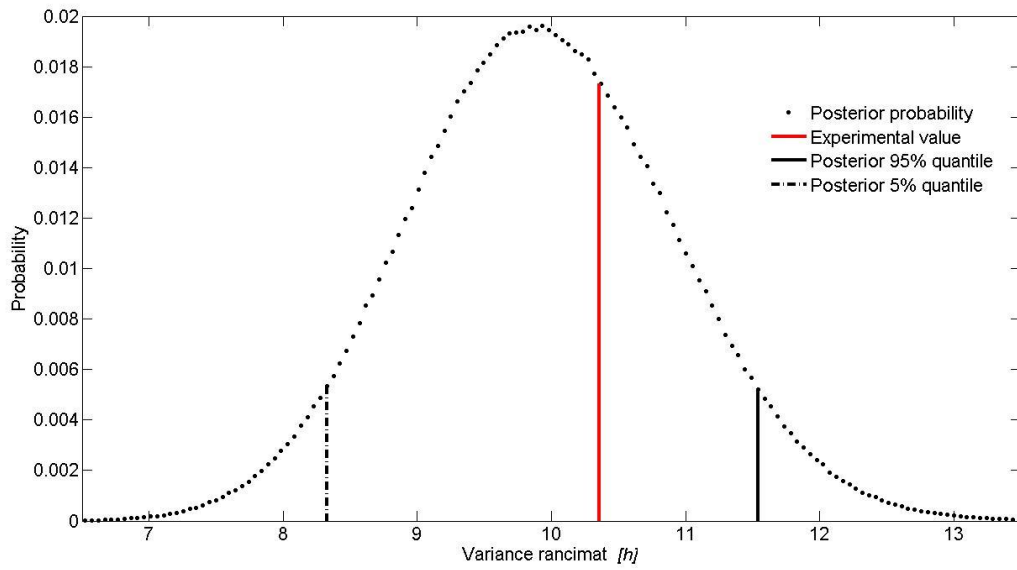
Figure H.131: Validation variance rancimat value model 1st grade
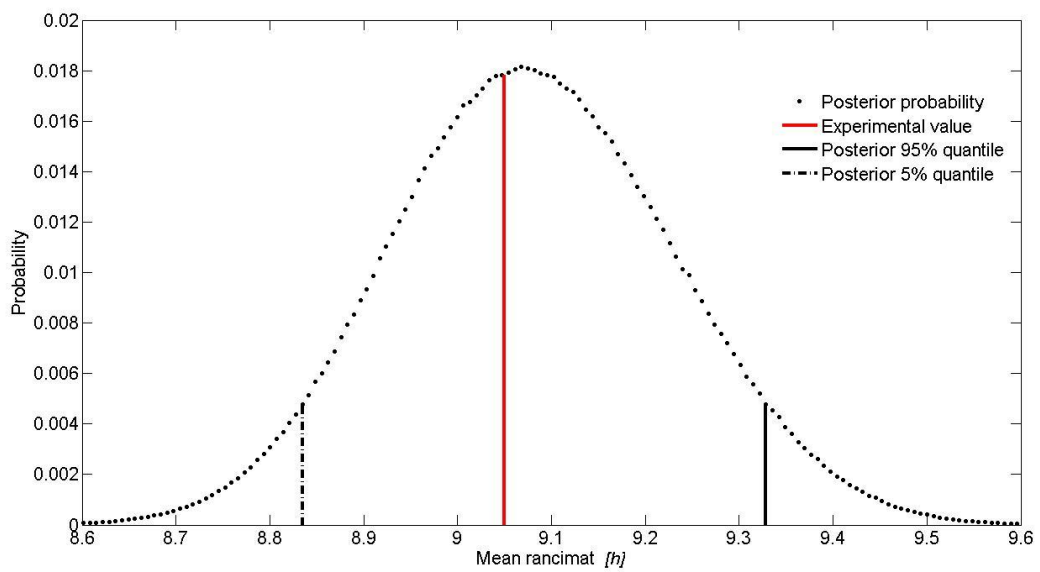
Figure H.132: Validation mean rancimat value model 2nd grade
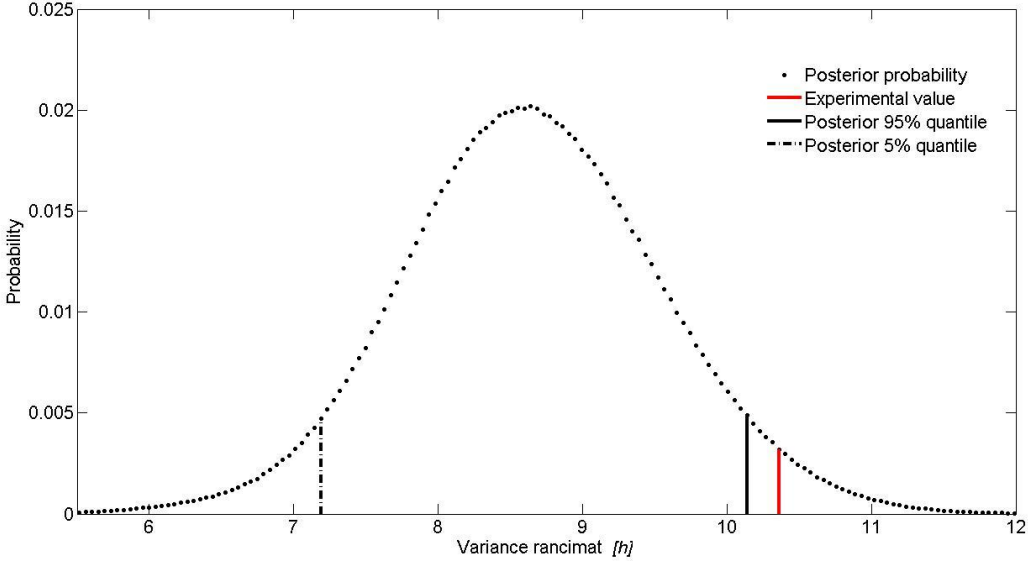
Figure H.133: Validation variance rancimat value model 2nd grade

Annex I – Figures for 5.2.2.2

    o   Concergence of Markov-Chains (exemplary)
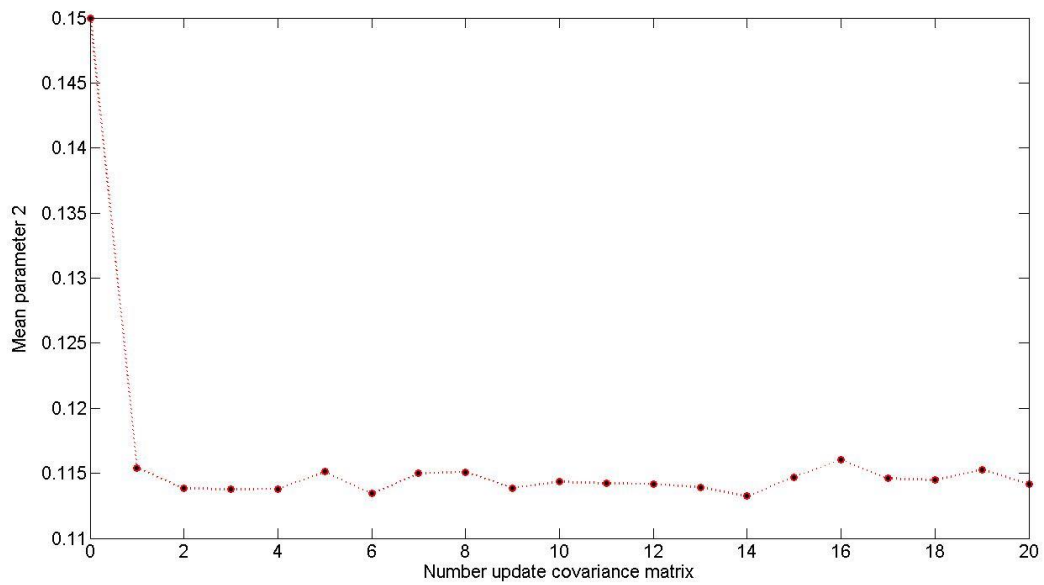
Figure I.134: Convergence mean parameter 2 model 1st grade acid value
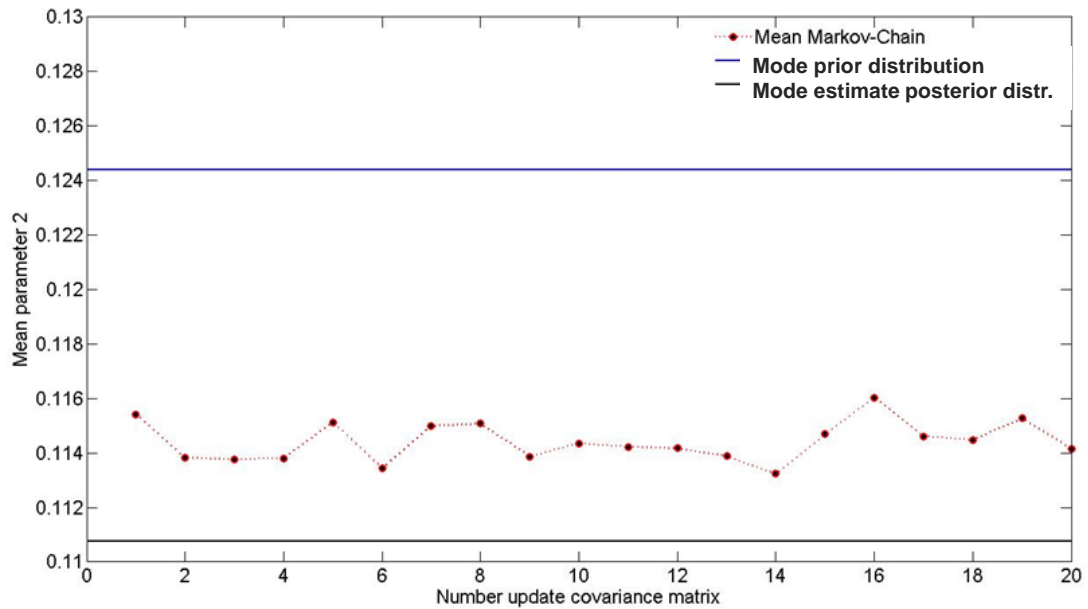
Figure I.135: Convergence mean parameter 2 model 1st grade acid value (zoom)
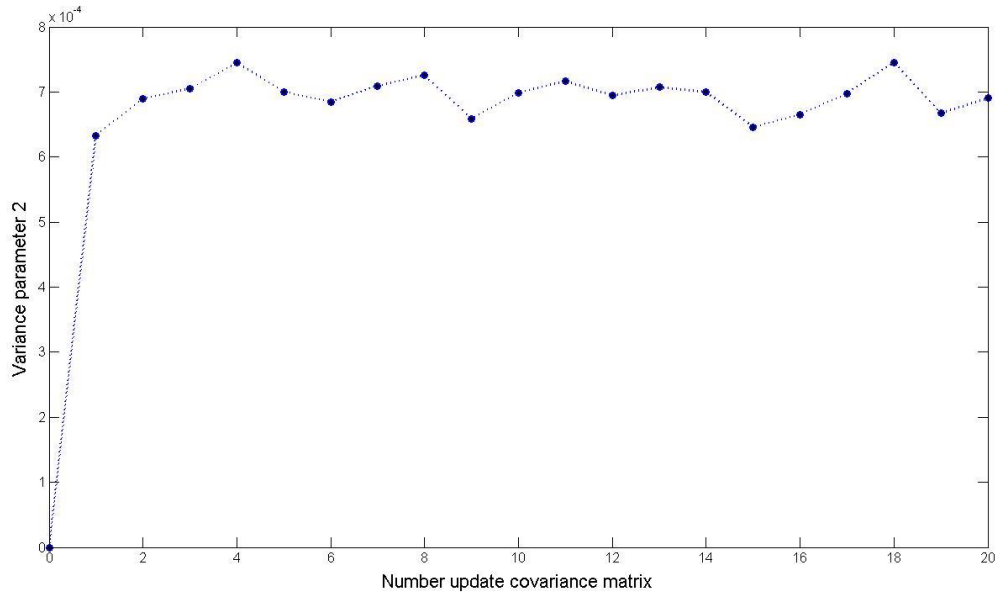
Figure I.136: Convergence variance parameter 2 model 1st grade acid value
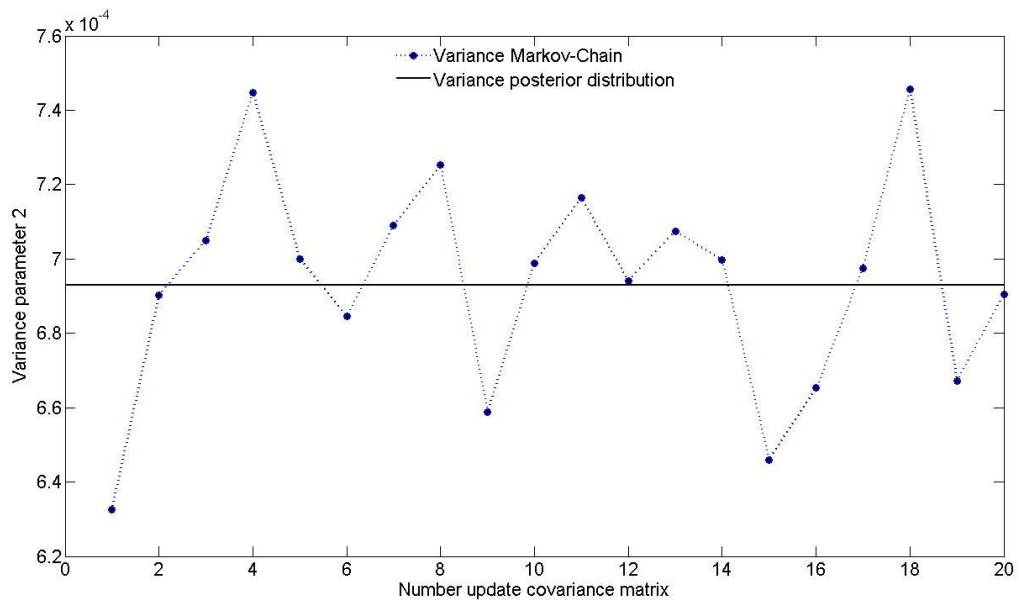
Figure I.137: Convergence variance parameter 2 model 1st acid value (zoom)

o Bayesian model fit
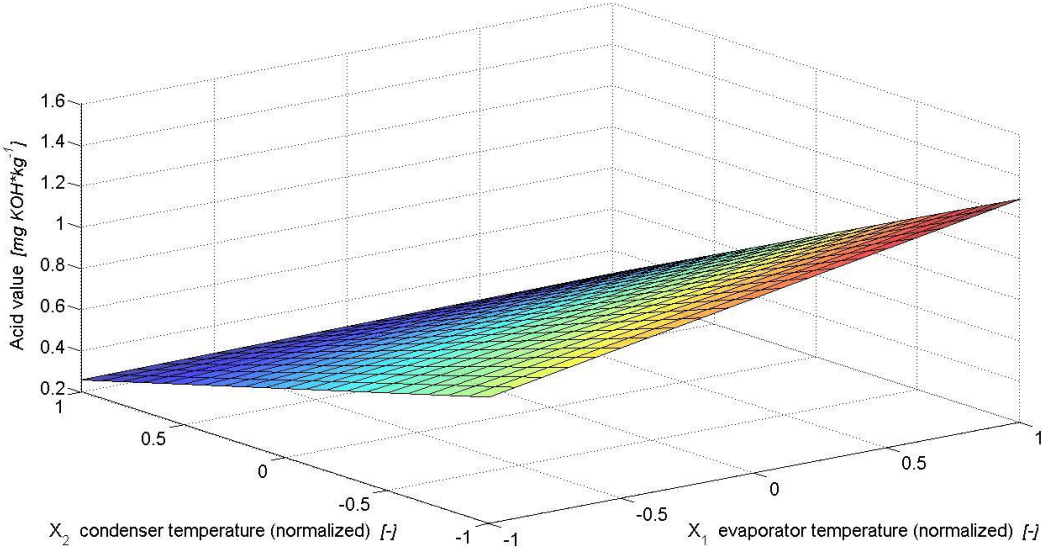
Figure I.138: Model 1st grade acid value for optimum pump power (Baysian fit)
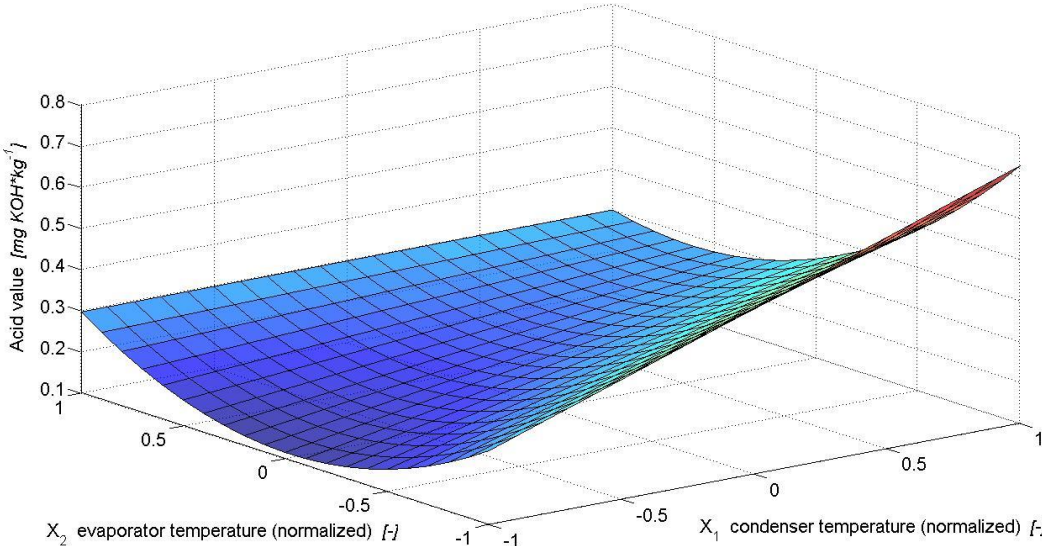
Figure I.139: Model 2nd grade acid value for optimum pump power (Baysian fit)

o   Prior and posterior parameter distribution (exemplary)

Figure I.140: Prior and posterior probability distribution of parameter 6 model 1st grade acid value

Figure I.141: Prior and posterior probability distribution of parameter 7 model 2nd grade acid value

o   Posterior probability distribution of experimental observations (exemplary)

Figure I.142: Posterior probability of exp. observations at opt. setting model 1st grade acid value

Figure I.143: Posterior probability of exp. observations at opt. setting model 2nd grade acid value
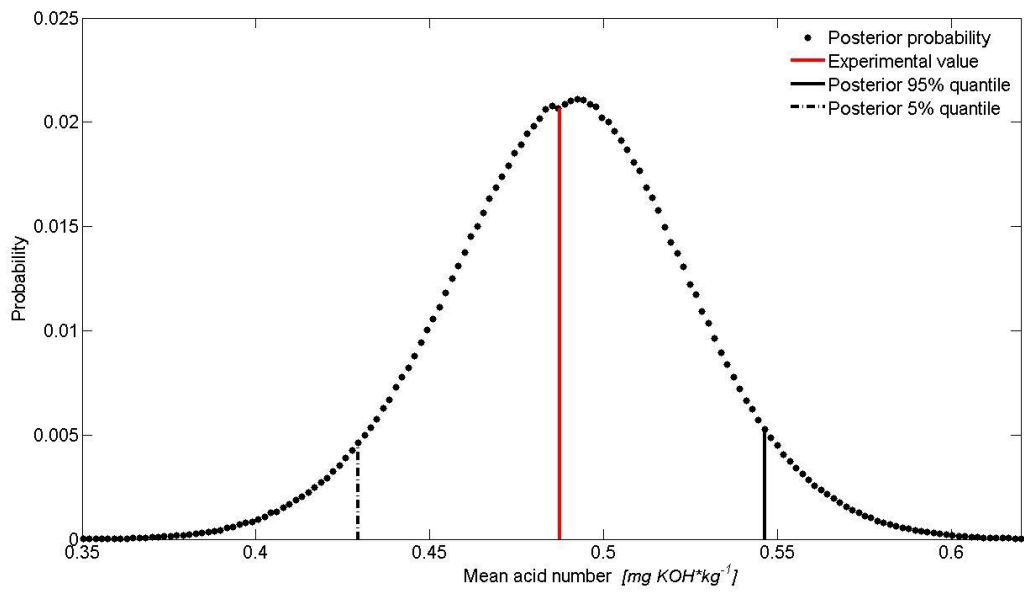
o Model validation

Figure I.144: Validation mean acid value model 1st grade
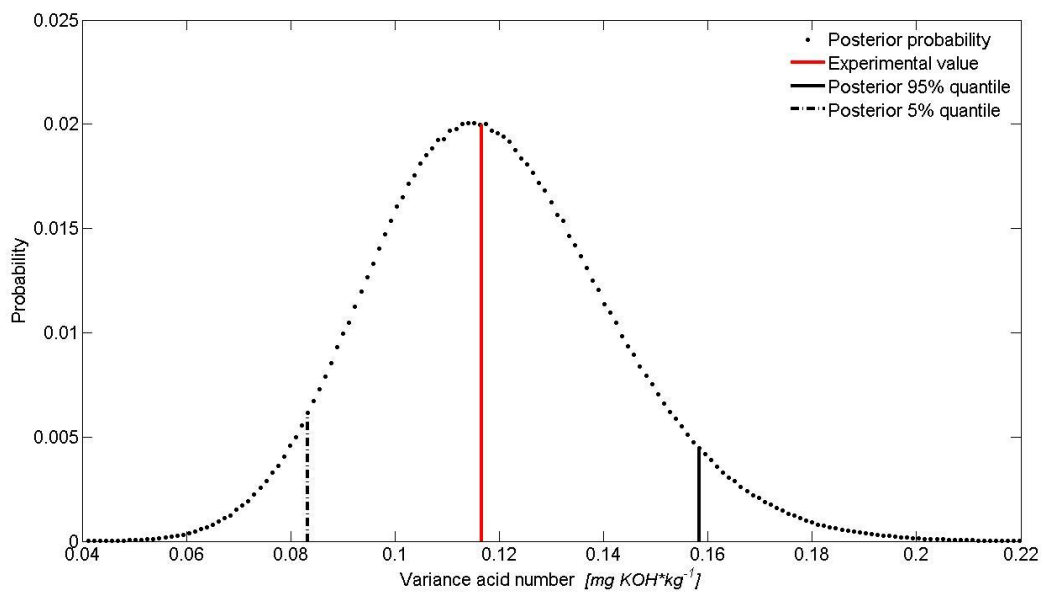
Figure I.145: Validation variance acid value model 1st grade
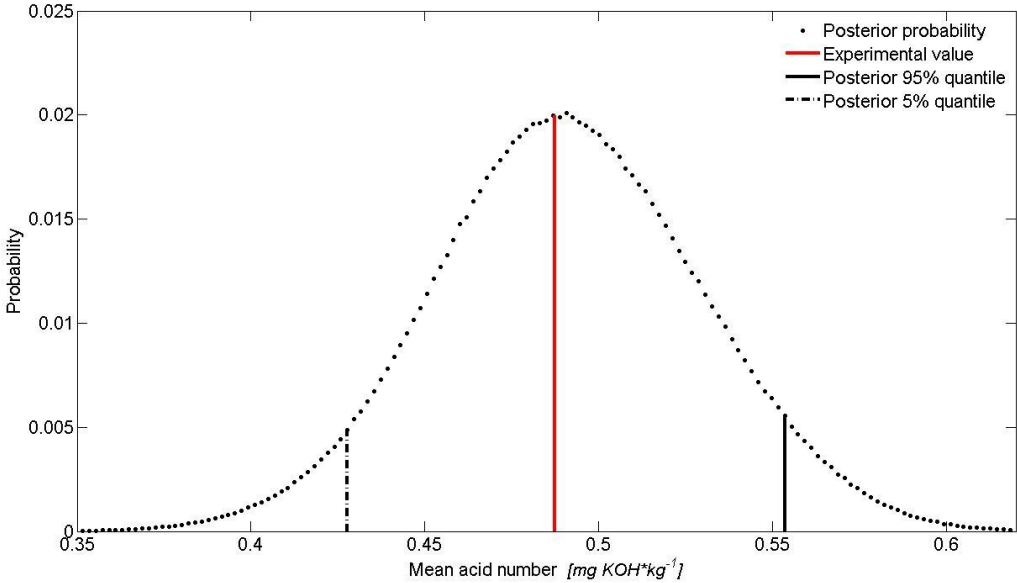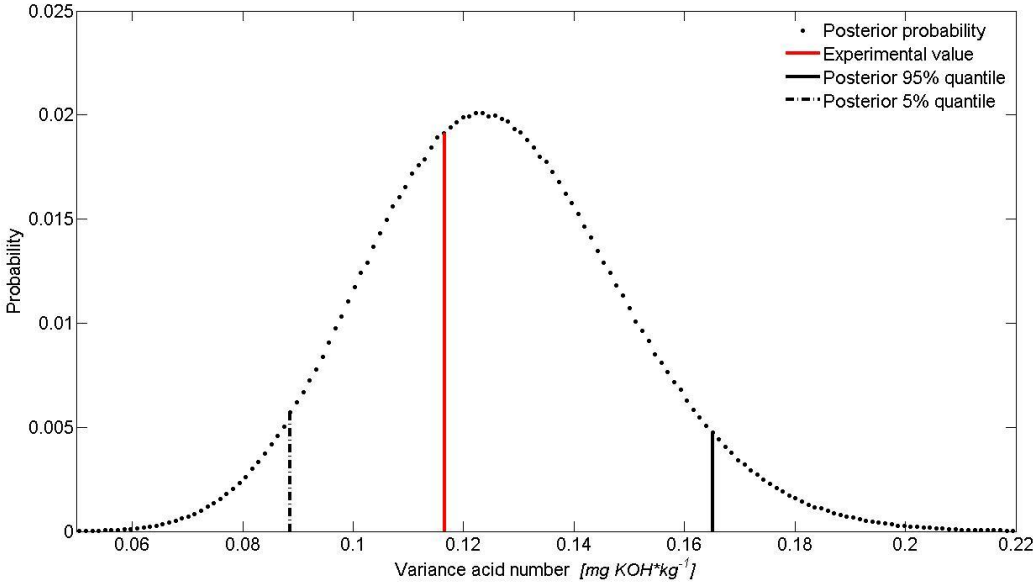
Figure I.146: Validation mean acid value model 2nd grade

Figure I.147: Validation variance acid value model 2nd grade

Annex J – Figures for 5.2.2.3
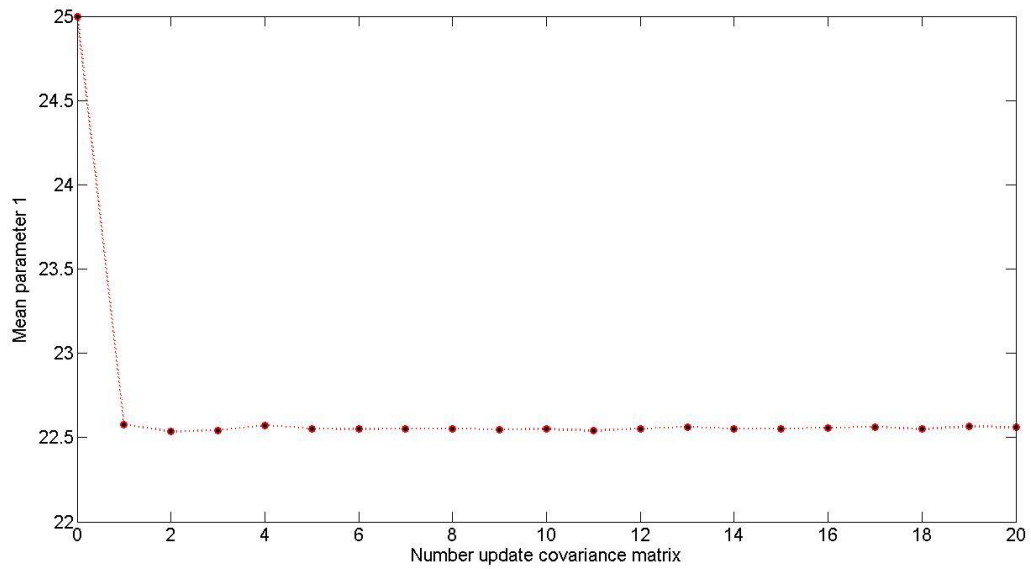
o Convergence of Markov-Chains (exemplary)

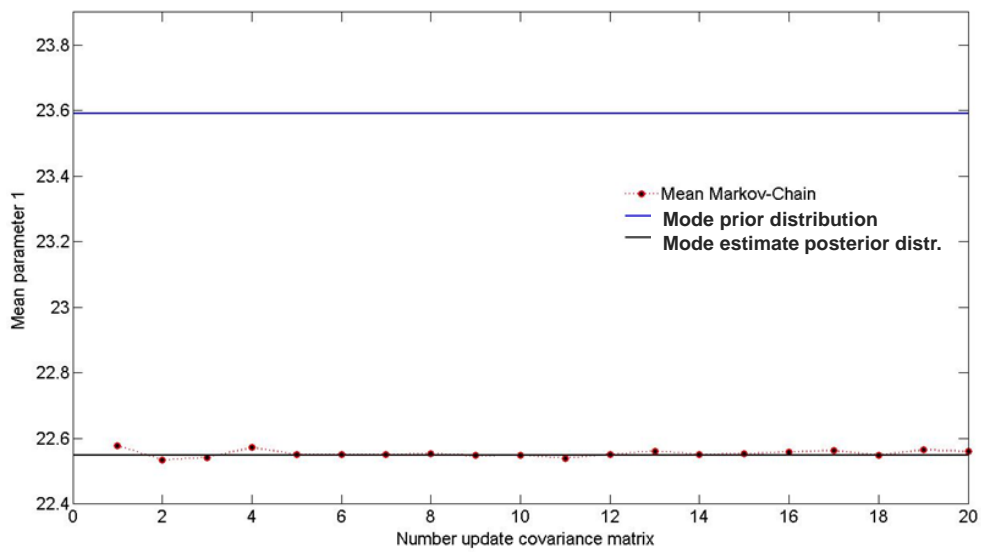Figure J.148: Convergence mean parameter 1 model 1st grade tocopherol

Figure J.149: Convergence mean parameter 1 model 1st grade tocopherol (zoom)
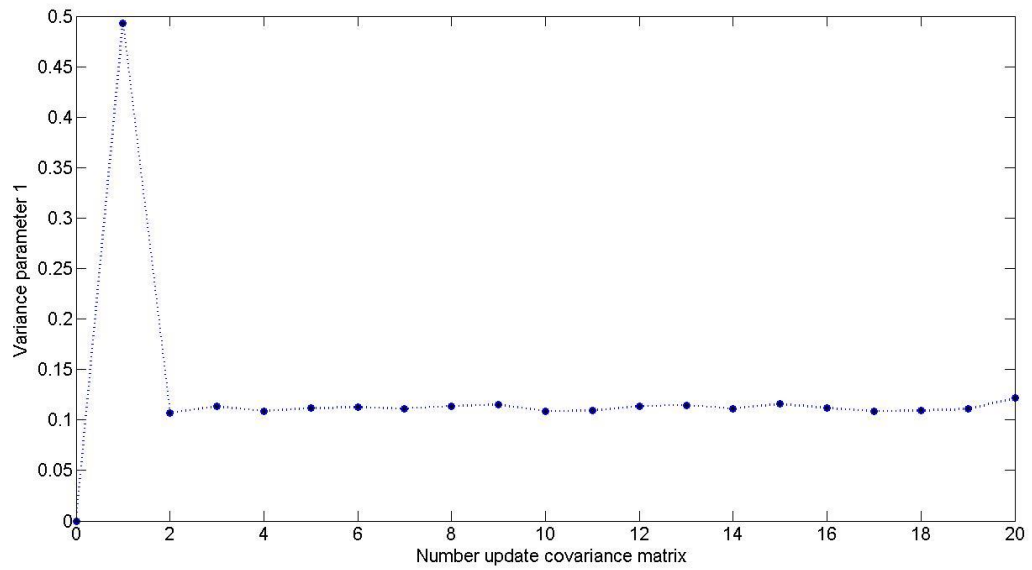
215

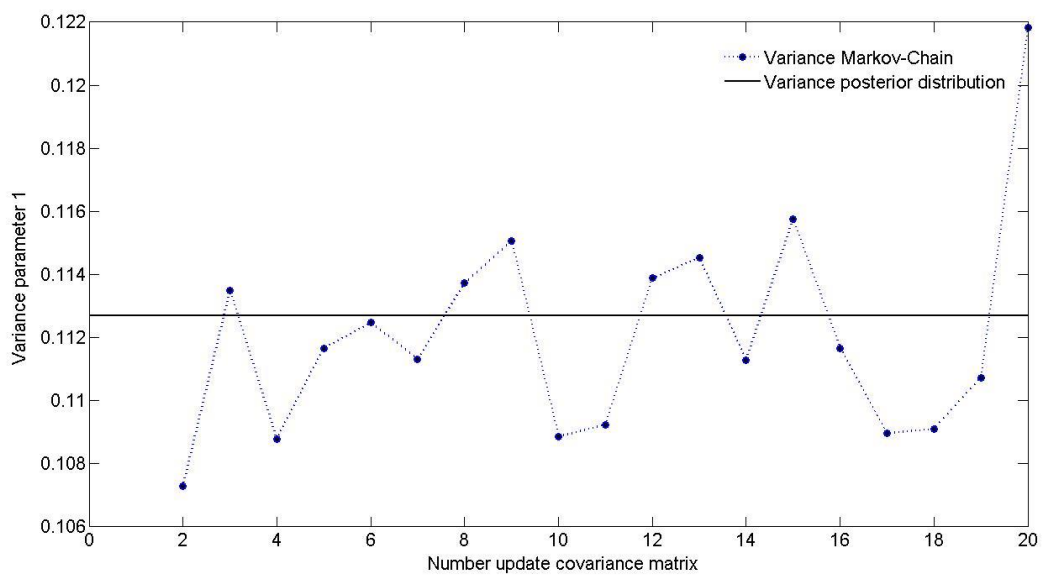Figure J.150: Convergence variance parameter 1 model 1st grade tocopherol

Figure J.151: Convergence variance parameter 1 model 1st grade tocopherol (zoom)
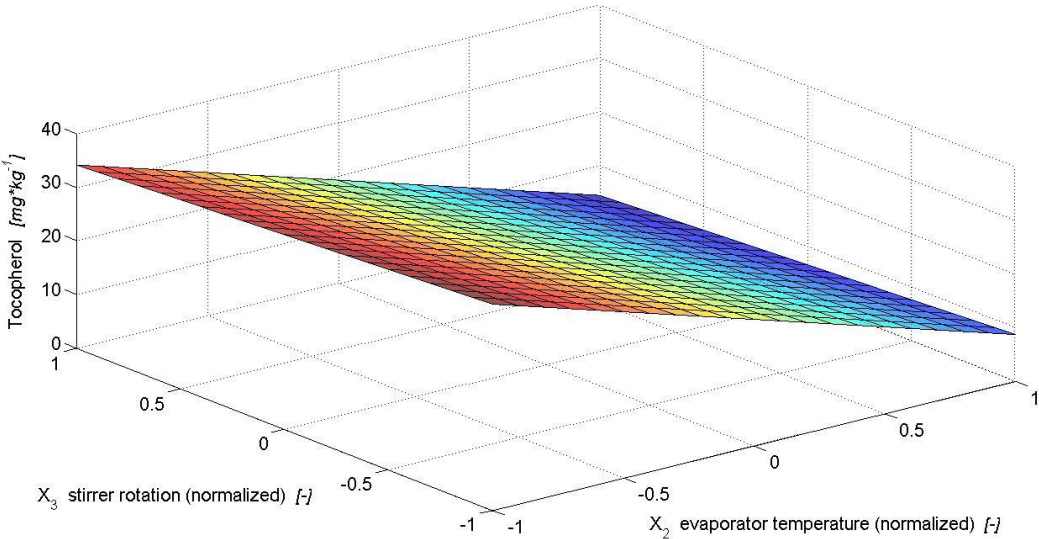
o   Bayesian model fit

Figure J.152: Model 1st grade tocopherol for optimum pump power (Baysian fit)
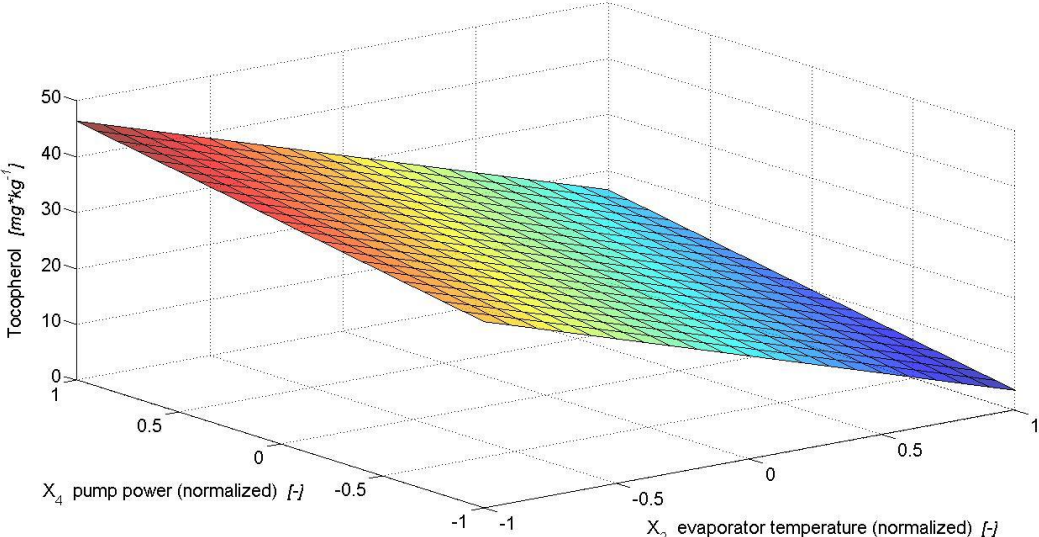

Figure J.153: Model 2nd grade tocopherol for optimum stirrer rotation (Baysian fit)

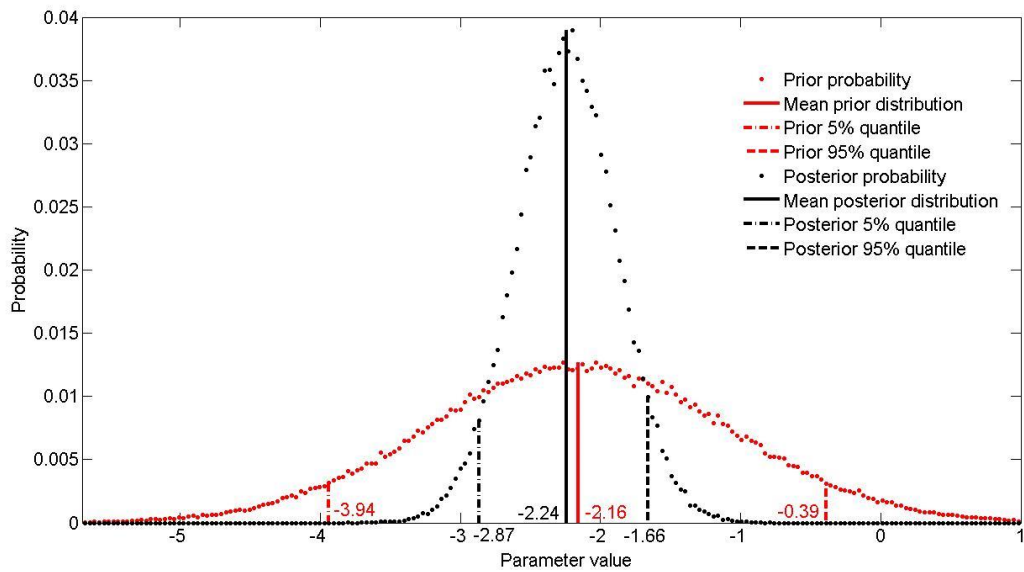o Prior and posterior parameter distribution (exemplary)

Figure J.154: Prior and posterior probability distribution of parameter 3 model 1st grade tocopherol
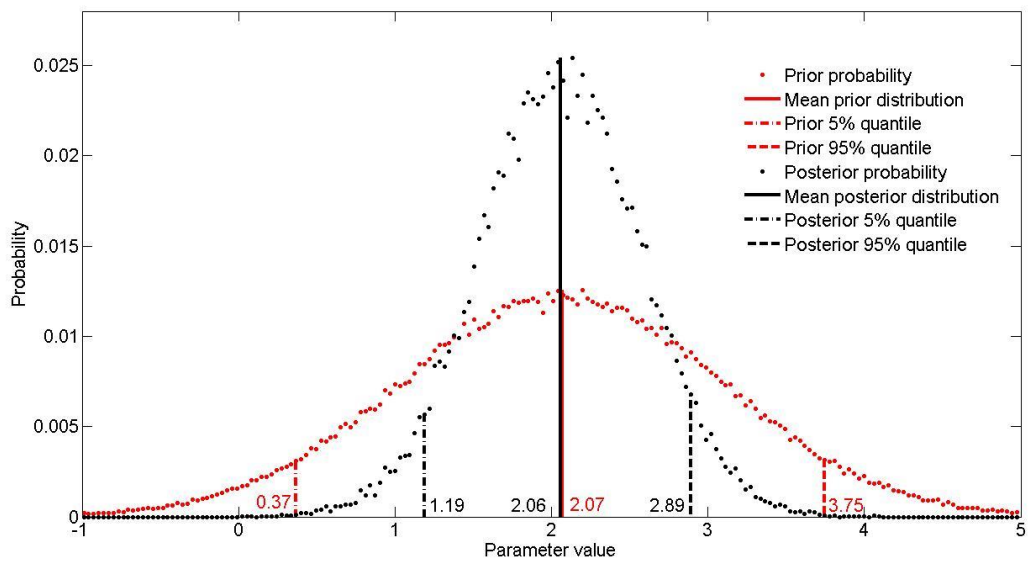

Figure J.155: Prior and posterior probability distribution of parameter 5 model 2nd grade tocopherol

o   Posterior probability distribution of experimental observations (exemplary)
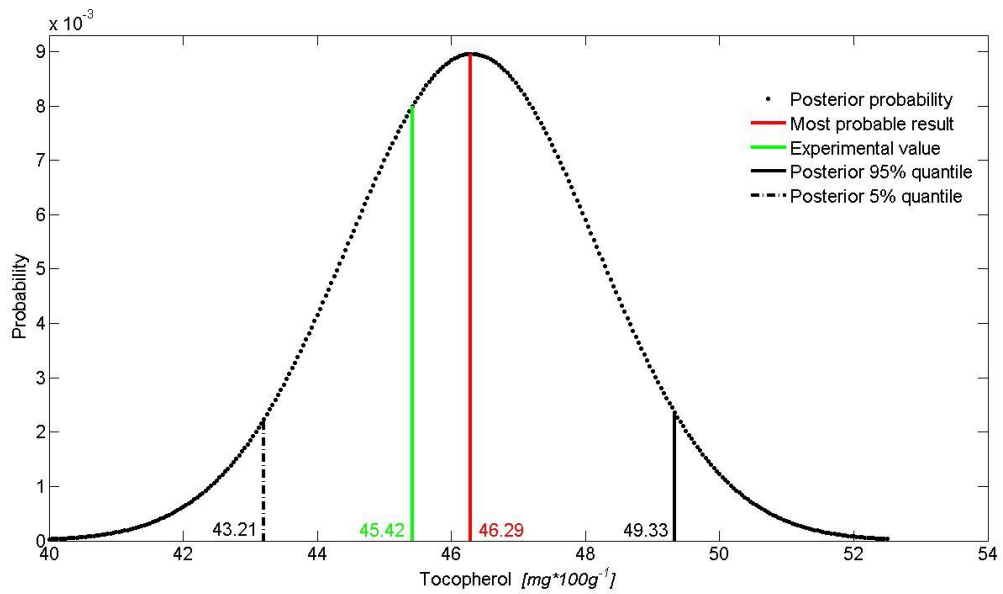
Figure J.156: Posterior probability of exp. observations at opt. setting model 1st grade tocopherol
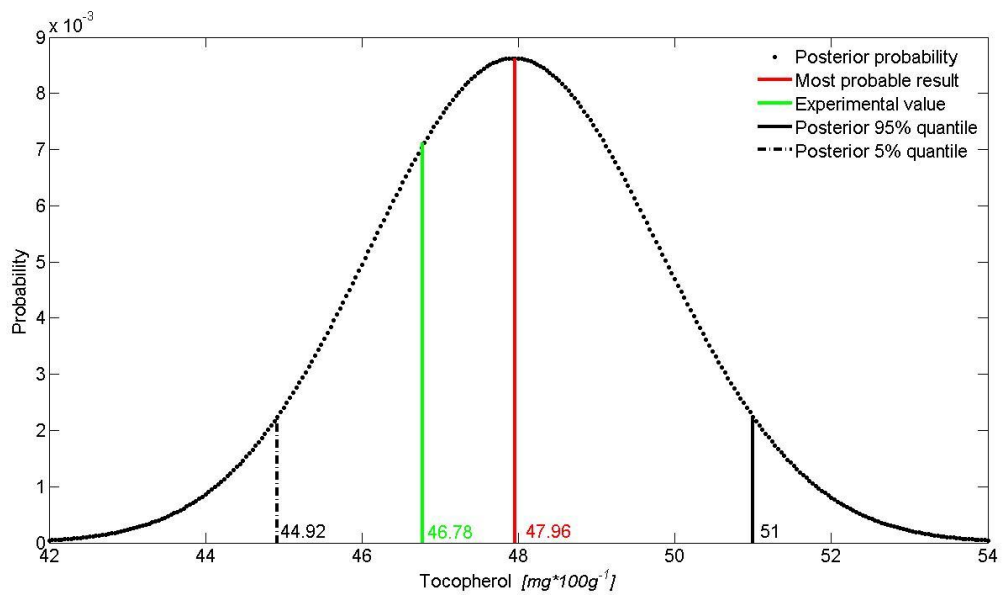

Figure J.157: Posterior probability of exp. observations at opt. setting model 2nd grade tocopherol
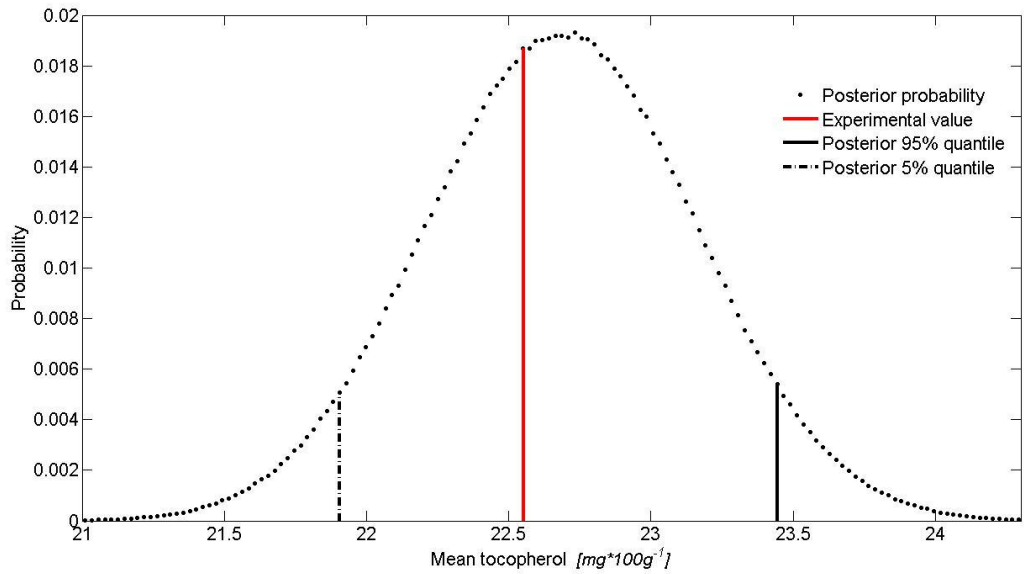
o   Model validation

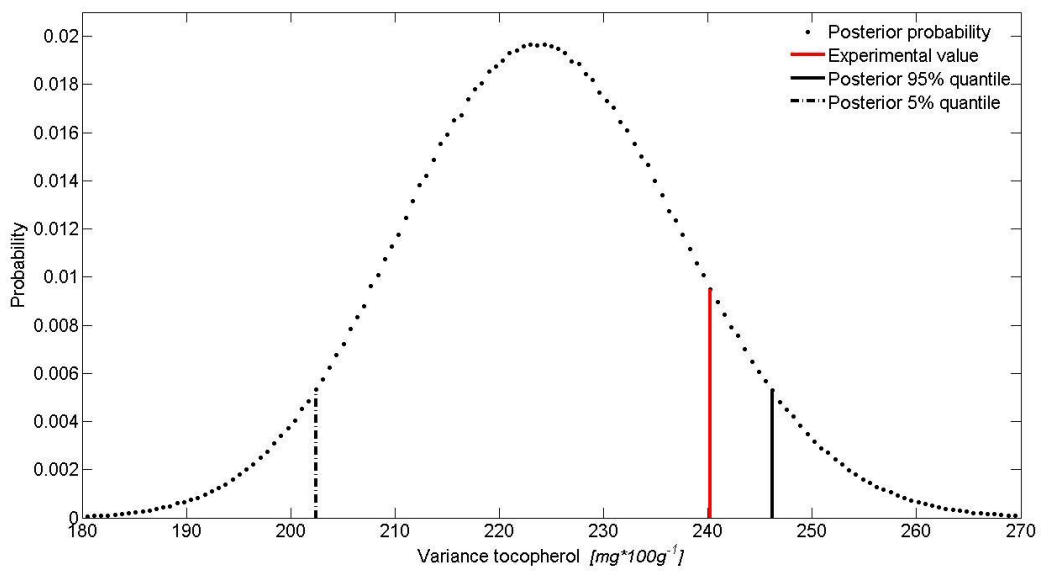Figure J.158: Validation mean tocopherol model 1st grade

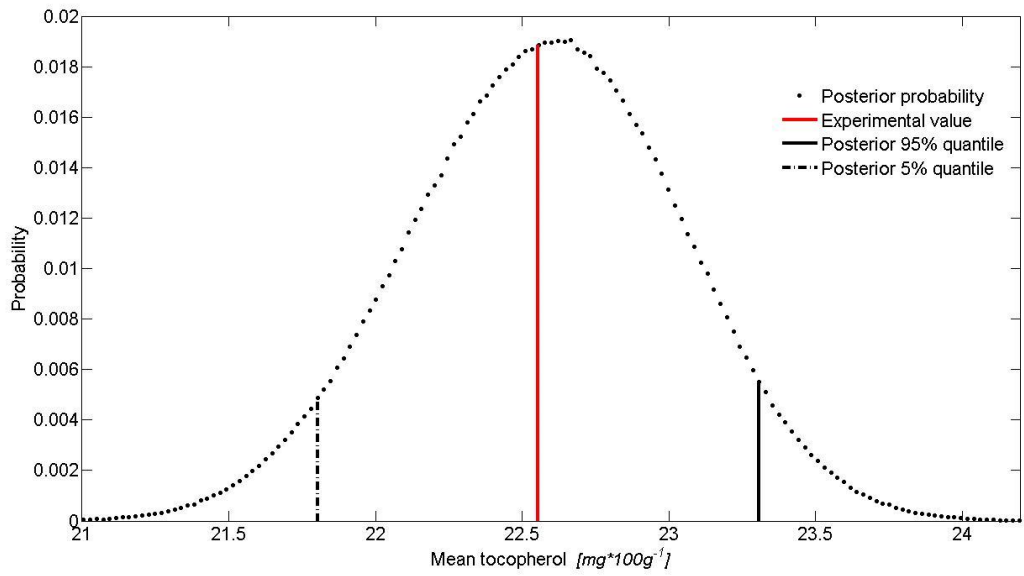Figure J.159: Validation variance tocopherol model 1st grade
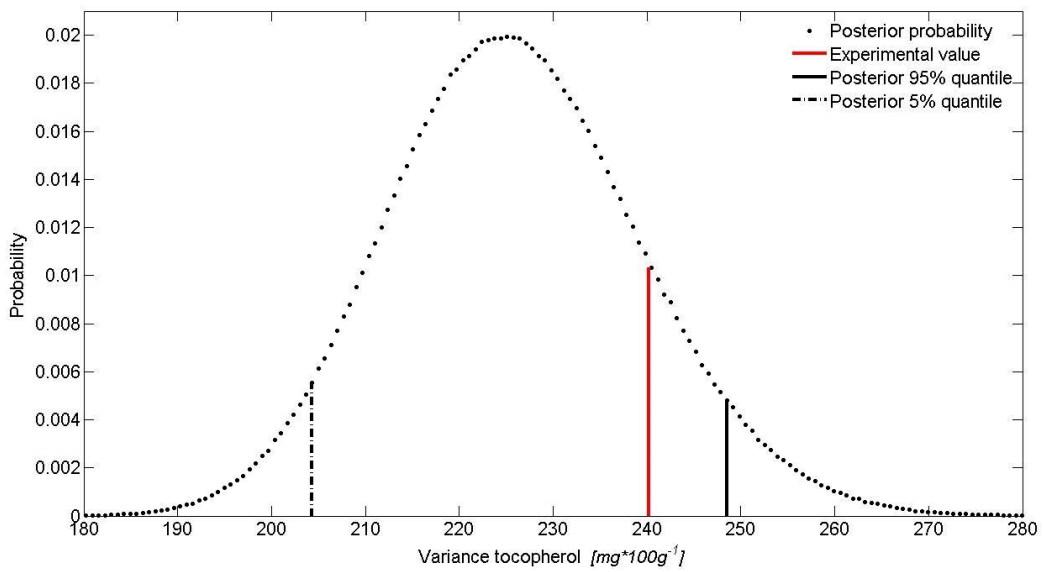
Figure J.160: Validation mean tocopherol model 2nd grade

Figure J.161: Validation variance tocopherol model 2nd grade

Bibliography

Abraham, Klaus; Appel, Klaus E.; Berger-Preiss, Edith; Apel, Elisabeth; Gerling, Susanne; Mielke, Hans et al. (2013): Relative oral bioavailability of 3-MCPD from 3-MCPD fatty acid esters in rats. In: *Arch Toxicol* 87 (4), pp. 649–659.

AIF-Project AIF 17059 BG „Grundlagen für die großtechnische Anwendung von Verfahren zur Herstellung von Speisefetten und -ölen mit reduzierten Gehalten an 3-MCPD-Fettsäureestern und verwandten Verbindungen" (Identification of novel technological strategies for minimizing the formation of carcinogenic substances during refining of edible plant oils), Final project report.

Amundsen, N.R., 1966, Mathematical Methods in Chemical Engeneering, Matrices and their Application. Prentice Hall, Engelwood Cliffs, N.J.

Andreß, H. J. (2003): F-Verteilung. Universität Bielefeld, online: http://eswf.uni-koeln.de/glossar/fvert.htm, last time checked 01.10.2015.

ASW, R. Wessendorf Agrar-Service u. Handel GmbH &. Co KG (Hg.) (2006): Spezifikation-Palmöl-Vollraffinat. Online: http://www.agrooil.de/Neue_Dateien/Spezifikation-Palmoel-Vollraffinat.pdf, last time checked 2015/1015.

Bakhiya, Nadiya, Abraham, Klaus, Gürtler, Rainer, Appel, Klaus Erich and Lampen, Alfonso, (2011), Toxicological assessment of 3-chloropropane-1,2-diol and glycidol fatty acid esters in food, in: Molecular Nutrition & Food Research, Volume 55, Issue 4, pp. 509–521.

Baltes, Werner; Matissek, Reinhard (2011): Lebensmittelchemie. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (Springer-Lehrbuch).

Bandemer, Hans; Bellmann, Andreas; Jung, Wolfhart; Richter Klaus (1973): Optimale Versuchsplanung. Berlin: Akademie Verlag (Wissenschaftliche Taschenbücher, Bd. 131).

Bethge, Daniel (1996): Destillation im Fein- und Hochvakuum. In: *Vakuum in Forschung und Praxis* 8 (2), pp. 84–87.

BfR (Hg.) (2007): Säuglingsanfangs- und Folgenahrung kann gesundheitlich bedenkliche 3-MCPD-Fettsäureester enthalten - Stellungnahme Nr. 047/2007 vom 11.12.07: Online: saeuglingsanfangs_und_folgenahrung_kann_gesundheitlich_bedenkliche_3_mcpd_fettsaeur eester_enthalten.pdf. Stellungnahme Nr. 047/2007 des BfR vom 11. Dezember 2007. Online: http://www.bfr.bund.de/cm/343/saeuglingsanfangs_und_folgenahrung_kann_gesundheitlich_ bedenkliche_3_mcpd_fettsaeureester_enthalten.pdf, zuletzt aktualisiert am 20.05.2013, zuletzt geprüft am 21.05.2013.

BfR (Hg.) (2009), Erste Einschätzung zur Bewertung der in raffinierten pflanzlichen Fetten nachgewiesenen Gehalte von Glycidol-Fettsäureester – Stellungnahme Nr. 007/2009), Online: erste_einschaetzung_von_glycidol_fettsaeureestern.pdf., last time checked: 15.10.2015.

Bhaggan, Krishnadath; Werleman; Jeanine Luvelle; Franx; Johan; 2011, Method of Treating a Vegetable Oil (US 20140357882 A1), in: http://www.google.com/patents/US20140357882

Bishop, D. J.; Nair, U. S. (1939): A Note on Certain Methods of Testing for the Homogeneity of a Set of Estimated Variances. In: *Supplement to the Journal of the Royal Statistical Society* 6, 1939 (1), pp. 89–99.

Bockisch, Michael (1993): Nahrungsfette und -öle. Stuttgart: Eugen Ulmer (Handbuch der Lebensmitteltechnologie).

Böttcher, Stefan; Stuttgart, CVUA (2007): CVUA Stuttgart | 3-MCPD-Ester in affinierten Speisefetten und Speiseölen - ein neu erkanntes, weltweites Problem. Untersuchungsämter-BW. Online: http://www.cvuas.de/pub/beitrag.asp?subid=1&Thema_ID=2&ID=717&Pdf=No, last time checked 15.10.2015.

Bibliography

Box, George E. P.; Draper, Norman Richard (2007): Response surfaces, mixtures, and ridge analyses. 2. Edition Hoboken, N.J: John Wiley (Wiley series in probability and statistics).

Box, George E. P. and Wilson, K. B., (1951), On the Experimental Attainment of Optimum Conditions, in: Journal of the Royal Statistical Society. Series B (Methodological), Vol. 13, No. 1 (1951), pp. 1-45.

BVE, Bundesvereinigung der Deutschen Ernährungsindustrie e.V. (Hg.), 2010, BVE – Fakten und Hintergründe zu nachhaltigem Palmöl. Online: www.bve-online.de/download/stelllungnahme-palmoel, last time checked: 21.05.2013.

Chib, Siddhartha and; Greenberg, Edward, 1995, Understanding the Metropolis-Hastings Algorithm, in: The American Statistician, Vol. 49, No. 4. (Nov., 1995), pp. 327-335.

Chib, S. and **Jeliazkov**, I. (2001), Marginal Likelihood from the Metropolis-Hastings-Output, in: Journal of the American Statistical Association, 96, pp. 270-281.

Congdon, Peter, 2006, Bayesian Statistical Modelling (2. Edition), John Wiley&Sons, Ltd.

Craft, B.D; Chiodini, A.; Garst, J.; Granvogl, M. (2013): Fatty acid esters of monochloropropanediol (MCPD) and glycidol in refined edible oils. In: *Food Additives & Contaminants: Part A* 30 (1), S. 46–51.

Craft, Brian D.; Nagy, Kornél; Sandoz, Laurence; Destaillats, Frédéric (2011): Factors impacting the formation of Monochloropropanediol (MCPD) fatty acid diesters during palm (Elaeis guineensis) oil production. In: *Food Additives & Contaminants: Part A*, pp. 1–8.

Craft, Brian D.; Nagy, Kornél; Seefelder, Walburga; Dubois, Mathieu; Destaillats, Frédéric (2012): Glycidyl esters in refined palm (Elaeis guineensis) oil and related fractions. Part II: Practical recommendations for effective mitigation. In: *Food Chemistry* 132 (1), pp. 73–79.

Creutzenberg, Otto, Berger-Preiß, Edith, 01/2011, 3-MCPD- und Glycidol-Fettsäureester - Stand zur Toxikologie und Untersuchungen zur Bioverfügbarkeit und Metabolisierung, Vortrag: http://www.ovid-verband.de/fileadmin/user_upload/ovid-verband.de/downloads/OVID_BLL_Informationsveranstaltung_Vortrag_Creutzenberg_Fraunhofer_ITEM.pdf.

Crews, C.; Brereton, P.; Davies, A. (2001): The effects of domestic cooking on the levels of 3-monochloropropanediol in foods. In: *Food Additives and Contaminants* 18 (4), pp. 271–280.

Crews, C.; Chiodini, A.; Granvogl, M.; Hamlet, C.; Hrnčiřík, K.; Kuhlmann, J. et al. (2013): Analytical approaches for MCPD esters and glycidyl esters in food and biological samples: a review and future perspectives. In: *Food Additives & Contaminants: Part A* 30 (1), pp. 11–45.

Crews, C.; Hough, P.; Brereton, P.; Harvey, D.; MacArthur, R.; Matthews, W. (2002): Survey of 3-monochloropropane-1,2-diol (3-MCPD) in selected food groups, 1999-2000. In: *Food Additives and Contaminants* 19 (1), pp. 22–27.

Davis, O. L. (1963), The Design and Analysis of Industrial Experiments, Oliver and Boyd, London

Destaillats, Frédéric; Craft, Brian D.; Dubois, Mathieu; Nagy, Kornél (2011): Glycidyl esters in refined palm (Elaeis guineensis) oil and related fractions. Part I: Formation mechanism. In: *Food Chemistry*.

Destaillats, Frédéric; Craft, Brian D.; Sandoz, Laurence; Nagy, Kornél (2012): Formation mechanisms of Monochloropropanediol (MCPD) fatty acid diesters in refined palm (Elaeis guineensis) oil and related fractions. In: *Food Additives & Contaminants: Part A* 29 (1), pp. 29–37.

EK, Die Kommission der europäischen Gemeinschaften (2006): Amtsblatt der Europäischen Union. Verordnung (EG) Nr. 1881/2006 der Kommission vom 19. Dezember 2006 zur

Festsetzung der Höchstgehalte für bestimmte Kontaminanten in Lebensmitteln, Online: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:364:0005:0024:DE:PDF.

FEI, Forschungskreis der Ernährungsindustrie e.V. (Hg.) (2011): „Industrielle Gemeinschaftsforschung: Instrument des innovativen Mittelstands". Dokumentation der FEI-Jahrestagung 2011. Strategien zur Minimierung von 3-MCPD-Fettsäureestern in Pflanzenölen. Unter Mitarbeit von B. Matthäus, pp. 47-73.

Frank, W. (1969): Die schonende Destillation Methoden und Erfahrungen bei der Destillation und Rektifikation empfindlicher Substanzen,. Mainz: Krausskopf-Verlag. Online verfügbar unter TP156.D5 F7.

Fregolente, L.V; Moraes, L.B; Martins, P.F; Batistella, C.B; Wolf Maciel, M.BR; Alfonso, A.P; Reis, M.H.M (2006): Distillation and Absorption 2006 (Symposium). Enrichment of natural products using an integratet solvent free process: molekular distillation: IChemE.

Gemeinschaftsausschuss für die Analytik von Fetten, Ölen Fettprodukten verwandten Stoffen und Rohstoffen (2012): Deutsche Einheitsmethoden zur Untersuchung von Fetten, Fettprodukten, Tensiden und verwandten Stoffen. Ergänzende Hinweise zu den DGF-Einheitsmethoden C-VI 17 (10) und C-VI 18 (10) zur Bestimmung der 3-MCPD- und Glycidyl-Ester. Online: http://www.dgfett.de/methods/hinweise.pdf, last time checked: 15.10.2015.

Geweke, J., 2007, Interpretation and inference in mixture models: Simple MCMC works. In Computational statistics and data Analysis, 5l(7), pp. 3529-3550.

Goffic, François Le; M. Albers (2002): Kurze Wege für große Moleküle. Short paths for big molecules. In: *Vakuum in Forschung und Praxis* 14 (5), pp. 292–295.

Habermeyer, Michael, Guth, Sabine and Eisenbrand, Gerhard, 03/2011, Identification of gaps in knowledge concerning toxicology of 3-MCPD and glycidol esters, in: European Journal of Lipid Science and Technology, Volume 113, Issue 3, pp. 314–318

Hamlet, C. G.; Jayaratne, S. M.; Matthews, W. (2002a): 3-Monochloropropane-1,2-diol (3-MCPD) in food ingredients from UK food producers and ingredient suppliers. In: *Food Additives and Contaminants* 19 (1), pp. 15–21.

Hamlet, C. G.; Sadd, P. A.; Crews, C.; Velíšek, J.; Baxter, D. E. (2002b): Occurrence of 3-chloro-propane-1,2-diol (3-MCPD) and related compounds in foods: a review. In: *Food Additives and Contaminants* 19 (7), pp. 619–631.

Hastings, W.K. (1970), Monte Carlo Sampling Methods using Markov Chains and their Applications, in Biometrika, pp. 57, 97-109.

Hartmann, K., Lezki, E., Schäfer, W. (1974), Statissche Versuchplanung und –auswertung in der Stoffwirtschaft, VEB Deutscher Verlag für Grundstoffindustrie, Leipzig.

Hickman, K. C. D. (1944): High-vacuum Short-path Distillation-A Review. In: *Chem. Rev.* 34 (1), pp.. 51–106.

Holló, J.; Kurucz, E.; Boródi, A. (1964): Versuche zur Entsäuerung von Sonnenblumenöl durch Molekulardestillation. In: *Fette, Seifen, Anstrichm.* 66 (11), pp. 936–941.

Hrncirik, Karel, 25/01/2010, Informationsveranstaltung: 3-MCPD- und Glycidyl-Fettsäureester in Lebensmitteln von BLL und OVID, Online: http://www.ovid-verband.de/fileadmin/user_upload/ovid-verband.de/downloads/Unilever_Hrncirik.pdf, last time checked: 15.10.2015.

Ito, Vanessa Mayumi; Batistella, César Benedito; Maciel, Maria Regina Wolf; Filho, Rubens Maciel (2007): Optimization of tocopherol concentration process from soybean oil deodorized distillate using response surface methodology. In: *Appl Biochem Biotechnol* 137-140 (1-12), pp. 885–896.

Junker, Patrick (2011): www.oekotest.de ÖKO-TEST Online Thema Versprechen und Mythen der Lebensmittelindustrie. Acrylamid & Co. ÖKO-TEST Verlag, Augsburg. Online:

http://www.oekotest.de/cgi/index.cgi?artnr=11011;gartnr=91;bernr=04;seite=07;co=;suche=3-mcpd)meist, last time checked: 15.10.2015.

Kendall, N.G. and Stuart, A. (1962), The advanced Theory of Statistics, Vol. 1 und 2, Griffin, London.

Klein, Bernd (2011): Versuchsplanung - DoE. Einführung in die Taguchi/Shainin-Methodik. 3. Aufl., Oldenbourg, München.

Kleppmann, Wilhelm (2008): Taschenbuch Versuchsplanung. Produkte und Prozesse optimieren ; [CD inside]. 5. Aufl. München, Wien: Hanser (Praxisreihe Qualitätswissen).

Krell, Erich (1976): Handbuch der Laboratoriumsdestillation. Mit einer Einführung in die Pilotdestillation. 3. Aufl., A. Hüthig, Heidelberg.

Kreyszig, E. (1972), Statistische Methoden und ihre Anwendungen, Vandenhoeck und Ruprecht, Göttingen.

Lin C, Wu EMY, Lee CN, Kuo SL (2011) Multivariate Statistical Factor and Cluster Analyses for Selecting Food Waste Optimal Recycling Methods. Environ. Eng. Sci. 28: pp. 349-356.

Linder (1969), Planen und Auswerten von Versuchen, Birkhäuser, Basel.

Masch, L.-W (1950): Die Molekulardestillation im Laboratorium und in der Technik. In: *Chemie Ing. Techn.* 22 (7), pp. 141–146.

Matissek, Reinhard; Steiner, Gabriele; Fischer, Markus (2010): Lebensmittelanalytik. 4. Aufl. Berlin, Heidelberg: Springer (Springer-Lehrbuch).

Matthäus, Bertrand; Pudel, Frank; Fehling, Peer; Vosmann, Klaus; Freudenstein, Anne (2011): Strategies for the reduction of 3-MCPD esters and related compounds in vegetable oils. In: *Eur. J. Lipid Sci. Technol.* 113 (3), pp. 380–386.

Metrohm, (2015), Application Bulletin 204/2 e. Oxidation stability of oils and fats - Rancimat method. Online: http://www.metrohm.de/Produkte/Katalog.html?identifier=AB-204&language=de&name=Oxidationsstabilit%C3%A4t+von+%C3%96len+und+Fetten+%E2%80%93+Rancimat-Methode, last time checked: 15.10.2015.

Müller, Peter; Sanso, Bruno; Delorio Maria; (2004); Optimal Bayesian Design by Inhomogenous Markov Chain Simulation. Journal of the American Statistical Association, 99(467): pp. 788-798.

Müller, Peter; (1991); A Generic Approach to Posterior Integration and Gibbs Sampling, Technical Report # 91-09, Department of Statistics, Purdue University.

Müller-Mulot, W., Rohrer, G. and Medeweth, R. (1976), European Journal of Lipid Science and Technology, Volume 78, Issue 7, pp. 257–262, Online: http://onlinelibrary.wiley.com/doi/10.1002/lipi.19760780701/abstract#fn1, last time checked: 15.10.2015

Nagy, K.; Sandoz, L.; Craft, B.D; Destaillats, F. (2011): Mass-defect filtering of isotope signatures to reveal the source of chlorinated palm oil contaminants. In: *Food Additives & Contaminants: Part A* 28 (11), pp. 1492–1500.

NIST/SEMATECH (Hg.) (2012): e-Handbook of Statistical Methods. 5.3.3.6. Response surface designs. Online: http://www.itl.nist.gov/div898/handbook/pri/section3/pri336.htm, last time checked: 15.10.2015.

O'Brien, Richard D. (2009): Fats and oils. Formulating and processing for applications. 3. Aufl. Boca Raton: CRC Press.

Ooi C.K.; Choo Y.M.; Ong A.S.H. (1992): Refining of edible oil. Veröffentlichungsnr: Malaysian Patent No. MY 104059A; Australian Patent No. AU-A-31084/89.

Ooi, C. K.; Choo, Y. M.; Yap, S. C.; Basiron, Y.; Ong, A. S. H. (1994): Recovery of carotenoids from palm oil. In: *J Am Oil Chem Soc* 71 (4), pp. 423–426.

Ooi, C.K; Choo, Y.M; Yap, S.C; Ma, A.N (1996): Refinig of red palm oil. In: *Elaeis* 8 (1), pp. 20–28.

TTZ OvGU, Technologie-Transfer-Zentrum der Otto-von-Guericke-Universität Magdeburg (Hg.): Forschungsportal Sachsen-Anhalt (2012): Innovations- und Transfernetzwerk Wissenschaft-Wirtschaft. Projekttitel: Grundlagen für die großtechnische Anwendung von Verfahren zur Herstellung von Speisefetten und -ölen mit reduzierten Gehalten an 3-MCPD-Fettsäureestern und verwandten Verbindungen. Online: http://www.forschung-sachsen-anhalt.de/index.php3?option=projektanzeige&lang=0&perform=&menu_link_active=&pid=16 516, last time checked: 2015/10/15.

OVID, Verband der ölsaatenverarbeitenden Industrie in Deutschland e. V. (Hg.): OVID :: Pflanzenöle. Daten & Grafiken. Online: http://www.ovid-verband.de/unsere-branche/daten-und-grafiken/pflanzenoele/, last time checked: 2015/10/15.

Pingris, A; Fontain, Je (1958): Abstracs. Vacuum distillation. In: J Am Oil Chem Soc 35 (10).

PPM e.V. (Hg.) (2011): Bestimmung der Säurezahl.

Pudel, Frank; Benecke, Patrick; Fehling, Peer; Freudenstein, Anne; Matthäus, Bertrand; Schwaf, Andrea (2011): On the necessity of edible oil refining and possible sources of 3-MCPD and glycidyl esters. In: *Eur. J. Lipid Sci. Technol.* 113 (3), pp. 368–373.

Rahn, Anja Katerina Karin; Yaylayan, Varoujan Antranik (2011): What do we know about the molecular mechanism of 3-MCPD ester formation? In: *Eur. J. Lipid Sci. Technol.* 113 (3), pp. 323–329.

Ramli, Muhamad Roddy; Siew, Wai Lin; Ibrahim, Nuzul Amri; Hussein, Rabeaah; Kuntom, Ainie; Abd. Razak, Raznim Arni; Nesaretnam, Kalanithi (2011): Effects of Degumming and Bleaching on 3-MCPD Esters Formation During Physical Refining. In: *J Am Oil Chem Soc* 88 (11), S. 1839–1844.

Rawlings, Herbert W. (1939): Molecular distillation of soybean and corn oils. In: *Oil Soap* 16 (12), S. 231–232.

Rimbach, Gerald; Möhring, Jennifer; Erbersdobler, Helmut F. (2010): Lebensmittel-Warenkunde für Einsteiger. Berlin, Heidelberg: Springer (Springer-Lehrbuch).

Rudolph, Tim (2013): RSM Short Path Distillation, 2013. MATLAB-Programm.

Scheffler, Eberhard (1986): Einführung in die Praxis der statistischen Versuchsplanung. 2. Aufl. Leipzig: Deutscher Verlag für Grundstoffindustrie.

Siebertz, Karl; Bebber, David Theo van; Hochkirchen, Thomas (2010): Statistische Versuchsplanung. Design of Experiments (DOE). 1. Aufl. Heidelberg, Dordrecht [u.a.]: Springer (VDI-Buch).

SGS (2014), Labor-Analyse von 3-MCPD-, 2-MCPD- und Glycidylestern in Lebensmitteln, Online: http://www.sgsgroup.de/de-DE/Local/Germany/News-and-Press-Releases/2014/08/Analyse-von-3-MCPD-2-MCPD-Glycidol-und-Glycidylester.aspx, last time checked: 15.10.2015.

Staby, Arne, (2014), Mechanistic Modeling: Does it Have a Future in Process Development?, Online: http://www.processdevelopmentforum.com/articles/mechanistic-modeling-does-it-have-a-future-in-process-development-2/, last time checked 15.10.2015.

Stahl, Bernhard (1991): Short-Path-Distillation. In: *Vakuum in Forschung und Praxis* 3 (2), pp. 134–137.

Strijowski, Ulf; Franke, Knut (2011), Surface adsorption may reduce 3-monochlorpropandiol-forming substances in vegetable oils (Lecture), Online: http://www.ovid-

verband.de/fileadmin/user_upload/ovid-verband.de/downloads/OVID_BLL_Informationsveranstaltung_Vortrag_Strijowski_DIL.pdf, last time checked: 15.10.2015.

Sunahara G., Perrin I., Marchesini M, 1993, Carcinogenicity Study on 3-Monochloropropane-1,2-diol (3-MCPD) Administered in Drinking Water to Fischer 344 Rats. Unpublished report No. RE-SR 93003 submitted to WHO by Nestec Ltd, Research & Development, Switzerland.

Sundmacher, Kai, Statistische Planung und Auswertung von Versuchen (Vorlesungsmanuskript), Lehrstuhl SVT, Otto-von-Guericke Universität, Magdeburg, 2007, pp.1221-1264.

Svejkovska, B.; Novotny, O.; Divinová, V.; Reblova, Z.; Dolezal, M.; Velíšek, J. (2004): Esters of 3-Chloropropane-1,2-Diol in Foodstuffs. In: *Czech Journal of Food Sciences* 22 (5), pp. 190–196.

Synowietz, Claudia (Hrsg.): Taschenbuch für Chemiker und Physiker. Band II: Organische Verbindungen, 4. Auflage (1983), Springer, Berlin.

Tierney, L. (1994), Markov Chains for Exploring Posterior Distributions, in: Annals of Statistics, 22, pp. 1701-1762.

UIC GmbH (Hg.) (2007), Thin Film Evaporators and Short Path Distillators. How they work • Applications • Customer specific trial distillations. Online: http://www.artema.net/uic/pdf_esempi_applic/prospetti/Applications_2007_en.pdf, last time checked: 15.10.2015.

USDA, United States Department of Agriculture, Foreign Agricultural Services (Hg.) (2010): Commodity Intelligence Report. Indonesia: Rising Global Demand Fuels Palm Oil Expansion. Online: http://www.pecad.fas.usda.gov/highlights/2010/10/Indonesia/, checked last time: 2015/10/15.

VTA, Verfahrenstechnische Anlagen GmbH (Hg.) (2010), Spezifikation 1-4848. Einstufige Kurzweg-Destillationsanlage für den Laborbetrieb.

Weißhaar, Rüdiger, 2008, 3-MCPD-Ester in raffinierten Speisefetten und Speiseölen, Onlien: http://www.cvuas.de/pub/beitrag.asp?subid=1&Thema_ID=2&ID=786, last time checked: 15.10.2015.

Wittka, F. (1940), Neuere Methoden der präparativen organischen Chemie. 7. Molekulardestillation. In: *Angew. Chem.* 53 (49-50), pp. 557–567.

WHO, World Health Organization; Food and Agriculture Organization of the United Nations; ebrary, Inc (2007): Evaluation of certain food additives and contaminants. Sixty-seventh report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, Switzerland: World Health Organization (WHO technical report series, 940). Online: http://site.ebrary.com/lib/academiccompletetitles/home.action.

Zelinková, Z.; Svejkovská, B.; Velíšek, J.; Doležal, M. (2006): Fatty acid esters of 3-chloropropane-1,2-diol in edible oils. In: *Food Additives and Contaminants* 23 (12), pp. 1290–1298.

Zelinková, Zuzana; Doležal, Marek; Velíšek, Jan (2009): Occurrence of 3-chloropropane-1,2-diol fatty acid esters in infant and baby foods. In: *Eur Food Res Technol* 228 (4), pp. 571–578.

Zulkurnain, Musfirah; Lai, Oi Ming; Latip, Razam Abdul; Nehdi, Imededdine Arbi; Ling, Tau Chuan; Tan, Chin Ping (2012): The effects of physical refining on the formation of 3-monochloropropane-1,2-diol esters in relation to palm oil minor components. In: *Food Chemistry* 135 (2), pp. 799–805.