Improved methods in optimal design of experiments for determination of water absorption kinetics of cereal grains

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

Fakultät Naturwissenschaften
Universität Hohenheim

Institut für Lebensmittelwissenschaft und Biotechnologie
Fachgebiet Prozessanalytik und Getreidewissenschaften
Prof. Dr. Bernd Hitzmann

vorgelegt von
Olivier Paquet-Durand
Geboren am 15. Juli 1979 in Münster

Stuttgart 2016

Dekan:	Prof. Dr. rer. nat. Heinz Breer
1. berichtende Person:	Prof. Dr. rer. nat. Bernd Hitzmann
2. berichtende Person:	Prof. Dr. Ing. Reinhard Kohlus
Eingereicht am:	17. Juni 2016
Mündliche Prüfung am:	25. August 2016

Die vorliegende Arbeit wurde am 22. August 2016 von der Fakultät Naturwissenschaften der Universität Hohenheim als ,,Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften" angenommen.

Chapter 1

Preface

1.1 Acknowledgements

This thesis was created form february 2011 to june 2016 at the department of process analytics and cereal science of the university of Hohenheim. I want to show my gratitude to all the people who allowed me to write it.

First I want to thank my supervisor Prof. Dr. Bernd Hitzmann for his friendly and helpful assistance and support. Numerous suggestions and discussions helped me to significantly improve this work. I also want to thank my co supervisor Prof. Dr. Reinhard Kohlus, for his support. His more practical or engineering based point of view on certain topics provided some helpful insights.

Beyond that I want to acknowledge the following persons:

Dr. Bianca Grote for her de facto mentoring in the first year and for her efforts to maintain mental health of all the group members. Dr. Florian Hecker for sharing his insights in cereal technology and his baking skills. Annika Hitzemann for coffee breaks and interesting discussions about roller derby which turned out to be an interesting sport. Marc Stanke for helpful comments and discussions about programming in general and optimization paticularly. And for tank battles in the lunch break. Viktoria Zettel for her support in the water absorption measurements. And for tons of cake and cookies, most of it very tasty.

Furthermore I wish to express my gratitude for my actual as well as former colleagues. They all invested time and effort in fruitful and sometimes intensive discussions, proofreading or just coffee and cookies. They also provided a very nice working and living environment. Of course I also want to thank the rest of the staff of the process analytics and cereal science department.

My special thanks goes to my family. Without them and their long lasting support, this thesis would not have been possible.

1.2 Co-authors

This work has been partially published previously with the knowledge and approval of the supervisor Professor Dr. Bernd Hitzmann. The scientific work presented in this thesis was partially conducted in cooperation with co-authors from the university of Hohenheim.

Optimal experimental design for parameter estimation of the Peleg model
By Olivier Paquet-Durand, Viktoria Zettel and Bernd Hitzmann published in Chemometrics and Intelligent Laboratory Systems.
Viktoria Zettel assisted in the experimental part to characterize the model system and determine reasonable initial parameter estimates to start the optimal design procedure with.

Optimal design of experiments and measurements of the water absorption process of wheat grains using a modified Peleg model
By Olivier Paquet-Durand, Viktoria Zettel, Reinhard Kohlus and Bernd Hitzmann published in Journal of Food Engineering.
Viktoria Zettel assisted in the experimental part and the night time moisture measurements. Professor Dr. Kohlus co-supervised this publication.

Applying the bootstrap method for optimal design of experiments and compare it to the classical method based on Cramer-Rao lower bound
By Olivier Paquet-Durand, Viktoria Zettel and Bernd Hitzmann published in Journal of Chemometrics.
Viktoria Zettel assisted in the characterization of the model system which is basically the same in all three publications.

1.3 Publication list

Peer-reviewed publications

- O. Paquet-Durand, D. Solle, M. Schirmer, T. Becker, B. Hitzmann
  Monitoring baking processes of bread rolls by digital image analysis, Journal of Food Engineering, Volume 111, Issue 2, Pages 425–431, July 2012

- O. Paquet-Durand, V. Zettel, B. Hitzmann
  Optimal experimental design for parameter estimation of the Peleg model, Chemometrics and Intelligent Laboratory Systems, Volume 140, Pages 36-42, January 2015

- O. Paquet-Durand, V. Zettel, R. Kohlus, B. Hitzmann
  Optimal Design of experiments and measurement of the water absorption process of wheat grains using a modified Peleg model, Journal of Food Engineering, Volume 165, Pages 166-171, June 2015

- O. Paquet-Durand, V. Zettel, B. Hitzmann
  Applying the bootstrap method for optimal design of experiments and compare it to the classical method based on Cramer-Rao lower bound, Journal of Chemometrics, In Press, DOI:10.1002/cem.2816, 2016

- J. Bolling, A. Fröhling, J. Durek, O. Schlüter, O. Paquet-Durand, B. Hitzmann
  Kontaminationsmonitoring in der Fleischzerlegung mittels Autofluoreszenzsignaturen, FleischWirtschaft 93 (3), Pages 114–117, March 2013

- F. Paquet-Durand, A. Sahaboglu, J. Dietter, O. Paquet-Durand, B. Hitzmann, M. Ueffing, PA. Ekström
  How long does a photoreceptor cell take to die? Implications for the causative cell death mechanisms. Advances in Experimental Medicine and Biology. Volume 801, Pages 575-581, 2014

- A. Sahaboglu, O. Paquet-Durand, J. Dietter, K. Dengler, S. Bernhard-Kurz, PA. Ekström, B. Hitzmann, M. Ueffing, F. Paquet-Durand
  Retinitis pigmentosa: rapid neurodegeneration is governed by slow cell death mechanisms, Cell Death and Disease, e488, 2013

- F. T. Hecker, W. B. Hussein, O. Paquet-Durand, M. A. Hussein, T. Becker
  A case study on using evolutionary algorithms to optimize bakery production planning, Expert Systems with Applications, Volume 40, Issue 17, Pages 6837-6847, December 2013

- V. Zettel, M. H. Ahmad, A. Hitzemann, M. Nache, O. Paquet-Durand, T. Schöck, F. Hecker, B. Hitzmann
  Optische Prozessanalysatoren für die Lebensmittelindustrie, Chemie Ingenieur Technik, Volume 88, Issue 6, Pages 684-687, May 2016

- V. Zettel, O. Paquet-Durand, F. Hecker, B. Hitzmann
  Image Analysis and Mathematical Modelling for the Supervision of the Dough Fermentation Process. In: Book of Abstracts of the 19th ESAFORM Conference, Nantes (France), April 27-29th 2016

Oral Presentations

- M. Schirmer & O. Paquet-Durand, D. Solle, M. Hussein, M. Jekle, B. Hitzmann, T. Becker
  Der intelligente Backofen, AGF Tagung für Bäckereitechnologie, Detmold, 2010

- O. Paquet-Durand, B. Grote, B. Hitzmann
  2D-Fluoreszenzspektroskopie mit multivariater Datenanalyse, Dechema Weiterbildung, Prozessanalytik in der Biotechnologie, 2011

- O. Paquet-Durand, A. Hitzemann, C. Ranzan, B. Grote, B. Hitzmann
  Data Mining zur Optimierung spektroskopischer Analyseverfahren, Dechema Frühjahrstagung der Biotechnologen, Frankfurt, 2013

- O. Paquet-Durand, V. Zettel, B. Hitzmann
  Optimale Versuchsplanung für die Bestimmung der Wasseraufnahmekinetik von Getreidekörnern, Jahrestreffen der ProcessNet-Fachgruppe Lebensmittelverfahrenstechnik, Erlangen, 2016

- O. Paquet-Durand, V. Zettel, F. Hecker
  Vorteile der Versuchsplanung zur mathematischen Beschreibung der Wasseraufnahme von Getreide, Automatisierungstagung – Sensortechnik in der Lebensmittelbranche, Freising, 2016

- O. Paquet-Durand, V. Zettel, B. Hitzmann
  Optimal experimental design for the parameter estimation of hydration kinetics of wheat grains, Analytica conference, Munich, 2016

Poster Presentations

- O. Paquet-Durand, B. Grote, B. Hitzmann
  Fluoreszenzsensor FLURIMAT zum Monitoring von Sauerteigfermentationen, Dechema Jahrestagung der Biotechnologen, Aachen, 2010

- B. Hitzmann, M. Nache, O. Paquet-Durand, O. Schlüter, J. Bolling, H. Schmidt, R. Scheier
  Fluoreszenz - und Raman - Spektroskopie zum Monitoring der Fleischzerlegung, Biotechnica, Hannover, 2013

- O. Paquet-Durand, A. Hitzemann, B.Grote, B. Hitzmann
  Informationsgewinnung aus hochdimensionalen Daten mittels Ameisenkoloniealgorithmus, Dechema Infotag: "Von der Datenflut zum Erkenntnisgewinn- der richtige Umgang mit (Bio)-Prozessdaten", Frankfurt, 2013

- O. Paquet-Durand, M. Nache, B. Hitzmann, J. Bolling, O. Schlüter, R. Scheier, H. Schmidt
  Using Fluorescence and Raman spectroscopy to monitor meat processing, Sartorius PAT & QbD Forum, Göttingen, 2014

Other Publications

- R. Scheier, O. Paquet-Durand, H. Schmidt, B. Hitzmann
  Nicht-invasive pH-Messung mittels Ramanspektroskopie, VVD-Tagung 2012: Praxis trifft Wissenschaft, Tagungsband, Dresden, March 2012

1.4 Summary

Optimal design of experiments is a part of experimental design where a mathematical model of the process under consideration is required. The aim of an optimal experimental design is to find the optimal experiment setup to achieve a target objective. Optimal experiment setup could be for example ideal measurement points in space and time. A common target objective is for example a precise determination of the parameters or minimal parameter estimation errors of the mathematical model. This is usually achieved by using computer based numerical optimization methods.

In this thesis, the optimal design of experiments was applied to determine hydration kinetics of wheat grains. The tempering of cereal grains - the water absorption of for example wheat kernels before milling - is an important step, because the power consumption of roller mills depends on the moisture content of the cereal. The correct moisture content also simplifies the sifting of mill stocks and flour. Consequently a goal might be to determine an ideal measurement setup for an as accurate as possible parameter determination of the used mathematical models to describe the water absorption.

In the first study the used mathematical model was the Peleg model for which the optimal design of experiments was carried out while investigating how the optimization criterion will influence the result. The parameter estimation errors could be reduced by up to 62 % compared to a non-optimal equidistant experimental design. It has been shown that the individual parameter estimation errors vary significantly depending on the used criterion. In this application only the D-optimal experimental design can reduce the parameter estimation errors of both parameters. In case of the A, P_r and E criterion at least one of the two parameter error could be reduced significantly. As the numerical optimization is computationally demanding, an alternative method for the entire optimal experimental design was developed. This alternative method is based on a mathematical function which depends on the rough initial parameter values. This function allows optimal measuring points to be calculated directly and therefore much faster, than the usual optimal design approach using numerical optimization techniques. In case of the very commonly used D-optimality criterion, the derived function is the exact solution. The deviation of the parameter estimation errors acquired by using the approximative optimal design instead of a normal optimal design are mostly around 0.01 % and therefore negligible.

In the second study, the suitability of the Peleg model for water absorption kinetics of wheat grains was investigated closer. Cereal grains usually consist of three major components, bran layer, endosperm and germ. All these components have different water absorption kinetics. Therefore the normal two parameter Peleg model might be insufficient to describe the water

absorption process of cereal grains properly. To address this, the Peleg model was enhanced and a second Peleg like term was added to account for the two biggest fractions of the grain, namely the endosperm and the bran layer. Two experiments were carried out, an initial experiment to get rough parameter values and a second experiment, which was then optimally designed. The modified Peleg model had now four parameters and could be used to describe the hydration process of wheat grains much more accurate. Using the parameters calculated from the initial experiment the optimal measurement points where calculated in a way that the determination of the parameters of the modified Peleg model was as accurate as possible. The percentage parameter errors for the four parameters in the initial experiment were 669 %, 24 %, 12 %, and 2.4 %. By applying the optimal design, they were reduced to 38 % 5.4 %, 4.5 % and 1.9 % respectively. The modified Peleg model resulted in a very low root mean square error of prediction of 0.45 % where the normal Peleg model results in a prediction error of about 3 %.

In the third study, it was investigated if bootstrapping could be used as a feasible alternative method for optimal experimental design. The classical procedure to determine parameter estimation errors is based on the Cramér-Rao lower bound but bootstrapping or re-sampling can also be used for the estimation of parameter variances. The newly developed method is more computationally demanding compared to the Cramér-Rao lower bound approach. However bootstrapping is not bound to any restrictive assumptions about the measurement and parameter variations. An optimal experimental design based on the bootstrap method was calculated to determine optimal measurement times for the parameter estimation of the Peleg model. The Cramér-Rao based optimal design results were used as a benchmark. It was shown, that a bootstrap based optimal design of experiments yields similar optimal measurement points and therefore comparable results to the Cramér-Rao lower bound optimal design. The parameter estimation errors obtained from both optimal experimental design methods deviate on average by 1.5 %. It has also been shown, that the probability densities of the parameters are asymmetric and not at all normal distributions. Due to this asymmetry, the estimated parameter errors acquired by bootstrapping are in fact likely to be more accurate. So bootstrapping can in fact be used in an optimal design context. However, this comes at the cost of a high computational effort. The computation time for a bootstrap based optimal design was around 25 minutes compared to only 5 seconds when using the Cramér-Rao lower bound method. But compared to the time required to carry out the experiments this is neglectable. Furthermore as computers get faster and faster over time, the computational demand will become less relevant in future.

1.5 Zusammenfassung

Die optimale Versuchsplanung ist ein Spezialfall der Versuchsplanung, bei der ein mathematisches Modell zur Beschreibung des betrachteten Prozesses notwendige Vorraussetzung ist. Das Ziel der optimalen Versuchsplanung ist, einen idealen Versuchs- und Messaufbau zu bestimmen, um ein gegebenes Ziel, z.B. die Minimierung der Parameterschätzfehler des verwendeten mathematischen Modells, bestmöglich zu erreichen.

Gegenstand dieser Arbeit war die Versuchsplanung für die Bestimmung der Wasseraufnahmekinetik von Weizenkörnern. Das Tempern, also das gezielte Befeuchten, von Weizen vor dem Vermahlen ist ein entscheidender Prozessschritt, da sowohl der Energiebedarf beim Mahlen als auch die Effektivität des Siebens beim Fraktionieren und Abtrennen der Schalenbestandteile maßgeblich vom korrekten Feuchtegehalt der Körner abhängen. Daher ergibt sich das Ziel, durch optimale Versuchsplanung die Parameter der Wasseraufnahmekinetik der Weizenkörner möglichst genau zu bestimmen.

Als Modell für die Wasseraufnahmekinetik wurde das Peleg Modell verwendet. Für dieses wurde die Versuchsplanung zur Minimierung der Parameterschätzfehler durchgeführt und der Einfluss unterschiedlicher Optimalitätskriterien evaluiert. Im Gegensatz zu einem äquidistanten Versuchsplan konnten durch die optimalen Versuchspläne die Parameterfehler um bis zu 62 % reduziert werden. Es zeigte sich außerdem, dass sich die "optimalen" Parameterschätzfehler je nach Optimalitätskriterium teils signifikant unterschieden. Für das Peleg Modell führt nur eine D-optimale Versuchsplanung zu einer Verringerung beider Parameterschätzfehler. Im Falle der A-, P_r und E-Optimalität verringert sich nur einer der beiden Parameterschätzfehler signifikant. Da die verwendeten numerischen Optimierungsverfahren verhältnismäßig rechen- und zeitaufwändig sind, wurde anschließend als Alternative eine Näherungsfunktion entwickelt, mit der sich optimale Messzeitpunkte als Funktion der Parameterschätzwerte direkt berechnen lassen. Im Fall der D-Optimalität entspricht diese Näherung einer analytisch exakten Lösung. Für die anderen Optimalitätskriterien ergeben sich geringe Abweichungen zu den optimalen Messzeitpunkten, die durch klassische Optimierung erhalten wurden. Diese Abweichungen ergeben auf die Parameterschätzfehler bezogen eine Differenz von etwa 0,01 % und wurden daher als vernachlässigbar betrachtet.

In der zweiten Publikation wurde die Eignung des Peleg Modells zur Beschreibung der Wasseraufnahmekinetik von Weizenkörnern näher untersucht. Getreidekörner bestehen im Wesentlichen aus drei Komponenten: Schale, Endosperm und Keimling. Alle Bestandteile unterscheiden sich in ihrem Wasseraufnahmeverhalten. Daher kann das einfache zwei-Parameter Peleg Modell die Wasseraufnahme von Getreidekörnen nur unzureichend beschreiben. Um eine bessere Beschreibung der Wasseraufnahme zu erreichen

wurde das klassische Peleg Modell um einen weiteren Peleg-Term ergänzt. Damit kann die Wasseraufnahme Kinetik von Schale und Endosperm getrennt abgebildet werden. Der Keimling, der nur etwa 3 % der Gesamtmasse ausmacht, wird vernachlässigt. Zur Modellerstellung wurden zwei Versuche durchgeführt. Zunächst ein initialer Versuch, um die vier Parameterwerte grob zu schätzen, und anschließend ein optimal geplanter Versuch, für den die Parameterschätzwerte benötigt wurden. Die Parameterschätzfehler konnten durch die optimale Versuchsplanung von initial 669 %, 24 %, 12 %, und 2,4 % auf 38 % 5,4 %, 4,5 % und 1,9 % respektive reduziert werden. Mit dem so modifizierten Peleg Modell ist der mittlere Vorhersagefehler für die Wasseraufnahme von Weizenkörnern mit 0,45 % sehr klein, im Gegensatz zu den etwa 3 % für das zwei-Parameter Peleg Modell.

In der dritten Publikation wurde untersucht, ob auch das Bootstrapping Verfahren zur Parameterfehlerschätzung im Rahmen einer optimalen Versuchsplanung verwendet werden kann. Die klassische Vorgehensweise bei der optimalen Versuchsplanung zur Minimierung der Parameterschätzfehler basiert auf der Cramér-Rao unteren Grenze zur Abschätzung der Parameterfehler. Eine Alternative dazu ist Bootstrapping, mit dem die Parametervarianz ebenfalls abschätzt werden kann. Dieses Verfahren ist im Vergleich zum "Cramér-Rao untere Grenze" Ansatz sehr rechenaufwändig. Allerdings sind für das Bootstrapping keinerlei einschränkende Annahmen bezüglich der Verteilungen der Messwerte oder der Parameter notwendig. Also wurde die Versuchplanung einmal basierend auf Bootstrapping und einmal klassich durchgeführt und die Resultate wurden verglichen. Es konnte gezeigt werden, dass die erzielten Resultate mittels Bootstrapping und Cramér-Rao unterer Grenze vergleichbar sind. Die minimalen Parameterschätzfehler weichen um etwa 1,5 % voneinander ab. Im Rahmen dieser Untersuchung konnte auch gezeigt werden, dass die berechnete und offensichtlich asymmetrische Wahrscheinlichkeitsdichteverteilung der Parameter keinesfalls einer Normalverteilung entspricht. Daher kann vermutet werden, dass die Ergebnisse der Versuchsplanung mittels Bootstrapping zumindest in diesem Fall prinzipiell vertrauenswürdiger sind. Dieser Vorteil wurde allerdings "erkauft" mit einem höheren Rechenaufwand. Die benötigte Rechenzeit stieg von 5 Sekunden für die Versuchsplanung mit Cramér-Rao unterer Grenze auf etwa 25 Minuten für die Versuchsplanung mittels Bootstrapping. Verglichen mit der Dauer eines Experimentes sind diese 25 Minuten aber eher unbedeutend. Ferner werden Computer in absehbarer Zukunft vermutlich noch deutlich schneller werden.

Contents

Chapter 2

Introduction and outline

2.1 Introduction

In natural sciences the core of a theory is usually a mathematical model, for example a function, a differential equation or a system of differential equations.
Such mathematical model consists of two types of information. First of all, and most obvious is the entanglement or relation of its variables:

> *The force of gravity F_g of two objects is proportional to the inverse of the squared distance d and also proportional to the product of the two masses m_1 and m_2.*

$$F_g \propto \frac{m_1 \cdot m_2}{d^2}$$

The other type of information are the parameters, actual numerical values which turn the proportionalities into an exact relation:

> *The force of gravity of two masses equals the product of the two masses divided by squared distance multiplied by the gravitational constant G.*

$$F_g = G \cdot \frac{m_1 \cdot m_2}{d^2}, \qquad F_g = 6.674 \cdot 10^{-11} \tfrac{m^3}{kg \cdot s^2} \cdot \frac{m_1 \cdot m_2}{d^2}$$

The determination of the parameters of a mathematical model is as important as understanding the basic relations of the relevant variables. The knowledge of the actual parameter values turns a basic understanding of a system into the ability of an exact prediction of the system behavior. To acquire information about the basic relation of variables as well as exact parameter values in the context of theory crafting, the system of interest has to be characterized via observation or experimentation.

There are countless ways, clever as well as plain inappropriate, to conduct experiments. To avoid "inappropriate" and promote "clever" ways, a lot of thought has to be put in how to conduct an experiment before actually performing it.

- "What do I want to show, prove or disprove?"
- "How can I achieve this?"
- "What kind of experiment is required to do so?"
- "How do I need to setup the experiment and how should I perform it to be sufficient to achieve the set goals (with minimal effort)?"

This process of actually giving thought about how to physically conduct the experiment is usually referred to as design of experiments (DoE). This step is crucial for a correct scientific approach and it has been ever since. It is usually taken out intuitively by the experimenter based on his experience.

If the exact experiment and measurement setup (what to measure, where to measure, when to measure, ...) has to be determined there are some alternatives to "experience" of the experimenter. If for example the development of a time dependent Variable $f(t)$ has to be determined over time, a common and simple approach to this question would be equidistant measurements (Van den Bos, 2007).

> *Measure f every n seconds.*

When there is little a priori knowledge available about the rough development of $f(t)$, it would make sense to use it and measure the value of f more often, when it actually changes a lot, and less often, when it changes less. An example would be a peak or an asymptotic behavior in the signal of $f(t)$.

> *Measure f more often, when it changes and less often or not at all, when it does not change.*

If there is a mathematical model already available which can describe the time dependence of $f(t)$ sufficiently, an optimal design of the experiments could be applied to determine the parameters as accurate as possible, which is what this work is all about.

Optimal experimental design in general is the search for an ideal measurement setup or ideal measurement points in time for an optimal result. In this case the optimal result may be determined by minimal parameter estimation errors. But one has to be aware, that the optimal design can only be applied, if the mathematical model is suitable to describe the process in question.

> *Measure f at the ideal measurement times (and only then) to achieve minimal parameter estimation errors.*

2.1.1 Cramér-Rao's lower bound for parameter error estimation

These ideal measurement points have to be determined before the experiment. How can this be done? For the determination of parameter errors, the Fisher information has been used (Lindner and Hitzmann, 2006; Ataíde and Hitzmann, 2009). The Fisher information is a measure of the information carried by measurements regarding the parameters of a mathematical model which was used to describe the measurable variable (Fisher, 1935). If this information for a specific parameter is high, the measurements carry a lot of information and the parameter determination is rather accurate. If on the other hand, the information is low, the measurements do not carry lots of information to determine the parameter in question accurately.
The Fisher information $I(\vec{p})$ is calculated as follows:

$$I(\vec{p})_{i,j} = \sum_{k=1}^{N} \frac{1}{\sigma_k^2} \cdot \frac{\partial f(t_k, \vec{p})}{\partial p_i} \cdot \frac{\partial f(t_k, \vec{p})}{\partial p_j}$$

Here $I(\vec{p})_{i,j}$ is the i, j^{th} element of the Fisher information, σ_k^2 is the measurement variance of the k^{th} measurement and p_i and p_j are the i^{th} and j^{th} parameter in the parameter vector $\vec{p}$.

The Fisher information can be used to estimate the variance of parameters which have been acquired for example by fitting a mathematical model to real world measurements by applying the least squares method.

The parameter variance estimation has been formulated in the Cramér-Rao lower bound (Rao, 1945; Cramér, 1946) which states that the inverse of the Fisher information is the lowest bound for the variances and covariances of the parameters (under ideal conditions). So by calculating the Fisher information and then the inverse, one could get the parameter variance-covariance matrix. The parameter estimation errors are then calculated as the square roots of the diagonal elements of the parameter variance-covariance matrix.

$$\Delta p_i \geq \sqrt{[I(\vec{p})^{-1}]_{i,i}}$$

One has to consider however, that this is only an estimation of the parameter variance as it is derived from an inequality. This is only true, if the measurement and parameter distributions are ideal (Kay, 2013).

2.1.2 Bootstrapping for parameter error estimation

An alternative method to calculate estimated values for the parameter variances is the resampling based bootstrapping method (Efron and Tibshirani, 1998), which has also been used extensively (Arai *et al.*, 2009; Banks *et al.*, 2010; Srivastava and Rawlings, 2014). But it is, at least compared to the Cramér-Rao lower bound based method, computationally very demanding. Bootstrapping relies on generating sub-samples of the measurement data by randomly picking measurement points with replacement. For each of the sub-samples the parameters of the used model are calculated by least squares fitting. The measurement error will necessarily result in slight variations in the calculated parameter values. So by repeatedly fitting a model to the sub-samples, the estimated distributions of the parameter values can be acquired. From these empirical parameter distributions the second moments can be calculated, which are an estimate of the standard error of the parameters. This method does not require any assumption of normality, neither for the measurement data nor for the estimated parameter variation. The probability density functions of both can be asymmetrically and of literally any shape.

2.1.3 Minimization of parameter estimation errors

Whichever method is used to determine the parameter estimation errors, the next and crucial step of optimal experimental design is then applying an optimization method such as for example the downhill simplex method (Nelder and Mead, 1965). The target value of the optimization would be the minimization of the parameter errors.

If there is more than one parameter error to minimize, one has to choose a feasible optimality criterion to address the multi objective optimization problem. An optimization criterion is a way of optimizing towards multiple, maybe excluding, goals. Some common criteria for optimal experimental design are (Pukelsheim, 1993):

- D-optimality $\rightarrow$ Maximize the determinant of the fisher information.
- E-optimality $\rightarrow$ Maximize the smallest eigenvector of the fisher information.
- A-optimality $\rightarrow$ Minimize the trace of the inverted fisher information.

By applying a numerical optimization algorithm to minimize these parameter estimation errors or a derived optimization criterion while changing the measurement points, it is possible to determine the optimal measurement points in time. This optimal design is usually an iterative process. The optimization is carried out, an according experiment is conducted, new parameter values are calculated and this new parameter values are compared to the initial

parameter values. If they differ too much, the optimal design procedure has to be repeated with the newly acquired parameters as initial guesses.

In optimal design of experiments, almost all numerical optimization algorithms are to a certain degree time consuming as they multiply the time required for the parameter error estimation by a large amount. That is probably the main reason why bootstrapping, which is by itself rather time consuming, has not yet been used for optimal experimental design.

2.1.4 Peleg's model for water absorption kinetics

In cereal technology the water absorption of cereal grains is important for milling and malting. Detailed knowledge is required to achieve correct moisture levels on the grains for best results (Butcher and Stenvert, 1973; Montanuci *et al.*, 2014). The Peleg model (Peleg, 1988) is a very simple but sufficient and therefore commonly used model for the description of water absorption of cereal grains (Sopade *et al.*, 1992) and legumes (Piergiovanni, 2011; Turhan *et al.*, 2002).

$$f(t, \vec{p}) = m_0 + \frac{t}{p_1 + p_2 \cdot t}$$

Here f is the amount of absorbed water after the time t, $\vec{p}$ is the parameter vector of the model with the elements p_1, p_2 and m_0 which is the initial moisture content.

The investigated system in this work was the water absorption process of wheat grains. All moisture uptake measurements were relative to the initial moisture content so that m_0 is 0 (wet based) and the Peleg equation has only two parameters.

2.2 Outline

The common theme for all three publications is optimal design of experiments for the hydration kinetics of wheat grains. The mathematical model used is based on the Peleg equation or variations of it. The actual goal of this investigation was to develop improvements to the optimal design procedure. The first possible improvement lies within the optimization step. For the optimization, the most commonly used methods are iterative numerical approximations as the downhill simplex method (Nelder and Mead, 1965) or genetic algorithm (Fraser, 1957) for more complex problems. They are easy to use and implement, but the problem with these methods is, that they can be computationally demanding and therefore time consuming.

Optimization algorithms are so ubiquitous and easy to implement, that one often would not even consider looking for an exact analytic solution. This would be a significantly better approach as one could skip the optimization step entirely. The computation of an analytic solution can also be a very demanding task. But if an optimal design has to be applied multiple times using the same model but with different parameters, this calculation has to be performed only once. Of course there might exist no analytic solution. In this case, one could still try to calculate or estimate a "good enough" approximation.

The title of the first publication is **"Optimal experimental design for parameter estimation of the Peleg model"**. It is about an investigation, if there is an analytic solution or a good approximation for an analytic solution to skip the optimization step.

During the measurements for the first publication, it became clear, that the Peleg equation was not ideal to describe the hydration/moisturisation process of wheat grains. So for the second publication **"Optimal design of experiments and measurements of the water absorption process of wheat grains using a modified Peleg model"** a more complex model was applied which was derived from the Peleg equation by adding two Peleg like terms. The assumption was that there are at least two major water absorbing components in the wheat grains which individually could be described by a Peleg term.

To this four parameter Peleg like model, the optimal design of experiment procedure was applied followed by experiments according to the determined optimal measurement points, to achieve high quality parameter values. Also the inverse problem, not the parameter error but the projected error in the independent variable, was examined. This is an interesting point for the tempering process in grain mills (Klingler, 2010) where the moisturisation of the grains has to be applied for a specific amount of time before milling. The required moisturisation times and the accuracy of those predicted times must be known (Shewry, 2009).

Another potential problem in the approved optimal design of experiment approach based on the Cramér-Rao lower bound and Fisher information matrix lies in the requirements about the

probability density of the measurement noise and the parameter probability density, which both have to be at least a continuous function. But this might not always be true (Pyzde, 1995). In case of the moisture absorption of wheat grains, it became obvious, that the parameter values were in fact not normally distributed. In the publication **"A bootstrap based method for optimal design of experiments"**, the feasibility of bootstrapping as an alternative for the Cramér-Rao lower bound for parameter estimation error determination was investigated.

Bootstrapping implies not restrictions about measurement or parameter distributions and is therefore much more flexible, and, assuming large sample sizes, more accurate due to no false assumptions. The disadvantage are the massive computational requirements compared to the "classical" optimal experimental design approach.

Chapter 3

Publications

3.1 Optimal experimental design for parameter estimation of the Peleg model

By Olivier Paquet-Durand, Viktoria Zettel and Bernd Hitzmann.

Published 2015 in **Chemometrics and Intelligent Laboratory Systems**.

Volume 140, Pages 36 - 42

Chemometrics and Intelligent Laboratory Systems 140 (2015) 36–42

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

Optimal experimental design for parameter estimation of the Peleg model

O. Paquet-Durand *, V. Zettel, B. Hitzmann

Department of Process Analytics and Cereal Science, Institute of Food Science and Biotechnology, University of Hohenheim, Garbenstr. 23, 70599 Stuttgart, Germany

ARTICLE INFO

Article history:
Received 15 July 2014
Received in revised form 26 September 2014
Accepted 26 October 2014
Available online 3 November 2014

Keywords:
Optimal experimental design
Parameter estimation
Water absorption
Peleg model
Fisher information
Cramer–Rao lower bound

ABSTRACT

In this work an optimal experimental design is applied to determine the parameters of the Peleg model for water absorption kinetics as precisely as possible. The parameter estimation errors were calculated using equidistant measurements in time as well as equidistant measurements on a logarithmic time scale. They were then compared to the optimal design results calculated, using multiple optimality criteria, constant measurement errors and the error of the initial rough parameter estimates. It is demonstrated that the optimal experimental design is beneficial for the parameter error estimation for at least one of the two parameters of the Peleg model equation. The parameter estimation errors could be reduced by up to 62% compared to an equidistant experimental design. Furthermore an approximation function for the optimal design process is developed. Depending on the used optimality criterion, with this function optimal measuring points for the Peleg model can be calculated directly and therefore much faster than the optimization procedure for the optimal experimental design. In case of the very commonly used D-optimality criterion, this function is even an exact solution. The deviation of the parameter estimation errors from the approximation function are mostly around 0.01% and therefore negligible.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The design of experiments is carried out to get most of the information with the smallest amount of investment. In optimal experimental design a mathematical model must be available and the goal of the procedure is to determine the model parameters with high precision. In recent years the number of scientific publications dealing with "Optimal Experimental Design" is increasing exponentially with a growth rate of 0.1 $year^{-1}$ and a $R^2 = 0.92$ (with respect to the Scopus data base).

Water absorption and swelling kinetics are key properties of many substances and food ingredients. They define their physical and chemical behaviour as well as shelf life of many if not all food products. Therefore it is vitally important to know these properties or, at least, to know how to measure and model them precisely.

Milk powder for example is mostly dehydrated for extended shelf life and economically transport. Usually milk powders are used rehydrated so the duration of the hydration process of dairy powders is very important to know [1,2]. In cereal technology the wheat flour production is also dependent on hydration as conditioning at defined temperatures for defined times is necessary to obtain better milling results [3]. Hydration of barley grains is preceding germination and malting and was modelled by Montanuci et al. [4]. They achieved the best results at 25 °C when the hydration needs only 12 h to reach the desired moisture content of 40% and propose their model for other grains to generate savings in time and cost. Sopade [5] demonstrated the suitability of Peleg's equation in modelling water absorption during soaking of cereal grains. During manufacture of legumes soaking is the first step but it is a long process (12–24 h). Piergiovanni [6] studied common bean seeds and identified differences in the speed of hydration. They found out that the Peleg model adequately describes the water absorption of fast hydrating common beans. For chickpeas soaking is used to gelatinize the starch in the grain. The Peleg model can be used to describe water absorption of chickpea between 20 and 100 °C as described by Turhan et al. [7]. The authors could demonstrate that the Peleg capacity constant is changing with temperature. It increases or decreases with increasing temperature depending on the sample and the method of moisture content calculation used.

To analyse the water absorption using the Peleg model, the measurements have to be carried out so that the parameters of the model can be determined as precisely as possible. However, in the food science literature there has yet been little attention paid to the precision of parameter estimation [8]. But precision of parameter estimation is important in process simulation and optimization and therefore, an optimal experimental approach is of first choice. Here the information from the Fisher information matrix can be used to identify optimal measurement points.

Mannarswamy et al. [9] investigated the optimal experimental design for three common adsorption isotherm models: the 2-parameter Freundlich, the 2-parameter Langmuir, and the 3-parameter Langmuir isotherm model. As optimization criterion they used the D-criterion

* Corresponding author. Tel.: + 49 711 459 22751; fax: + 49 711 459 23259.
E-mail address: o.paquet-durand@uni-hohenheim.de (O. Paquet-Durand).

http://dx.doi.org/10.1016/j.chemolab.2014.10.006
0169-7439/© 2014 Elsevier B.V. All rights reserved.

(as explained below). They discuss the variation of the design with the assumed values of the model parameters and demonstrated that the general equivalence theorem was satisfied for all three models. Ataíde and Hitzmann [10] used the optimal experimental design to analyse the Michaelis–Menten kinetics in a stirred tank reactor. They gave error ranges for the rough initial parameters in which the experimental design is in favour compared to ordinary equidistant measurement points (equally spaced observations). Donckels et al. [11] used the so called anticipatory design, where the expected information content of a newly designed experiment is considered before the experiment is carried out. They investigated the enzyme kinetic of the conversion of glucose by glucokinase (EC: 2.7.1.2) and would like to discriminate nine different models for this reaction. They concluded that the anticipatory approach definitely makes sense to design discriminatory experiments.

A list of the most recent applications of model based optimal experimental design techniques in various fields (from chemical kinetics to biological modelling) is presented by Franceschini and Macchietto [12]. They present main contributions to model-based experimental design procedures in terms of novel criteria, mathematical formulations and numerical implementations.

In this contribution the procedure of optimal experimental design for the Peleg model is presented to determine water absorption kinetics.

2. Theory

The main goal of optimal experimental design is to determine measurement points for an experiment in a select way that the parameters of the corresponding model can be calculated as precisely as possible. However, for the calculation of the optimal experimental design based on mechanistic models, rough parameter values must be known.

2.1. Cramer–Rao lower bound

If the measurement errors are normally distributed, the parameter estimation errors variance can be calculated based on the Cramer–Rao-lower-bound.

$$\mathbf{var}(\boldsymbol{p_i}) \geq \left[\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)^{-1}\right]_{\boldsymbol{ii}} \tag{1}$$

The Cramer–Rao lower bound is a limit of the performance of the best unbiased estimator for the parameters, which can be reached under optimal conditions. It is used as estimation for the parameter error variance. The lower bound of the estimation error variance var(…) of the $\boldsymbol{i}^{\text{th}}$ component of parameter vector $\vec{\boldsymbol{p}}$ is determined by using the diagonal elements (marked by $\ldots_{\boldsymbol{ii}}$) of the inverse (defined by $^{-1}$) of the Fisher information matrix $\boldsymbol{I}$. The $\boldsymbol{ij}$-element of the Fisher information matrix is defined as:

$$\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)_{\boldsymbol{ij}} = -\boldsymbol{E}\left\langle \frac{\partial^2 \ln(\mathbf{pdf}(\boldsymbol{y}, \vec{\boldsymbol{p}}))}{\partial \boldsymbol{p_i} \cdot \partial \boldsymbol{p_j}} \right\rangle \tag{2}$$

where $\boldsymbol{E}\langle\ldots\rangle$ is the expected value of the derivative of the probability density function **pdf** (…) of the measurement variable $\boldsymbol{y}$ with respect to the model parameters $\boldsymbol{p_i}$ and $\boldsymbol{p_j}$. The model for the measurement variable is given as:

$$\boldsymbol{y} = \boldsymbol{f}\left(\boldsymbol{x}, \vec{\boldsymbol{p}}\right) + \boldsymbol{e} \tag{3}$$

$\boldsymbol{x}$ stands for the independent variables, e.g. the coordinate or time. The measurement error $\boldsymbol{e}$ is assumed to be normal distributed with a constant sigma. If the errors are input dependent, another optimal design approach has to be performed according to e.g. Boukouvalasa et al. [13].

The Fisher information matrix is then approximated as follows:

$$\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)_{\boldsymbol{ij}} = \frac{\mathbf{1}}{\boldsymbol{\sigma^2}} \sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{N}} \cdot \frac{\partial \boldsymbol{f}\left(\boldsymbol{x_k}, \vec{\boldsymbol{p}}\right)}{\partial \boldsymbol{p_i}} \cdot \frac{\partial \boldsymbol{f}\left(\boldsymbol{x_k}, \vec{\boldsymbol{p}}\right)}{\partial \boldsymbol{p_j}} \tag{4}$$

Here $\boldsymbol{N}$ is the number of measurements and $\boldsymbol{\sigma^2}$ is the variance of the measurement error at the measurement point $\boldsymbol{x_k}$. So, the change of the function $\boldsymbol{f}$ with respect to a change of the parameter at the measurement points $\boldsymbol{x_k}$ has an influence on the element of the Fisher information matrix. If the change is high, the information is high, if on the other hand the function does not depend on the parameter at the measurement point, no additional information is generated. The partial derivatives of the function $\boldsymbol{f}\left(\boldsymbol{x}, \vec{\boldsymbol{p}}\right)$ with respect to one parameter are often called sensitivities.

2.2. The Peleg model

For water absorption various process models have been developed (e.g. [14]). A very common model to describe water absorption processes in food depending on time is the Peleg model [15]. The Peleg model expresses the water content of a material with respect to time and is defined as:

$$\boldsymbol{f}_{\text{peleg}}(\boldsymbol{t}) = \boldsymbol{m_0} + \frac{\boldsymbol{t}}{\boldsymbol{p_1} + \boldsymbol{p_2} \cdot \boldsymbol{t}} \tag{5}$$

Here $\boldsymbol{f_{peleg}}(\boldsymbol{t})$ is the water content at the time $\boldsymbol{t}$, $\boldsymbol{p_1}$, $\boldsymbol{p_2}$ and $\boldsymbol{m_0}$ are the parameters of the model. $\boldsymbol{m_0}$ is the starting water content of the sample, $\boldsymbol{p_1}$ is the inverse of the initial water uptake rate and $\boldsymbol{p_2}$ determines the final water content as can be seen, if $\boldsymbol{t}$ goes to infinity:

$$\lim_{\boldsymbol{t}\to\infty} \boldsymbol{f}_{\text{peleg}}(\boldsymbol{t}) = \boldsymbol{m_0} + \frac{\mathbf{1}}{\boldsymbol{p_2}} \tag{6}$$

If the initial water content of the observed compound is known or if it had been measured in advance, it can be excluded from the experimental design. The parameter count is then reduced to two. If Eq. (5) is applied to Eq. (4), the result is the Fisher information matrix for the Peleg model:

$$\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right) = \frac{\mathbf{1}}{\boldsymbol{\sigma^2}} \cdot \begin{pmatrix} \sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{N}} \frac{\boldsymbol{t_k^2}}{(\boldsymbol{p_1} + \boldsymbol{p_2} \cdot \boldsymbol{t_k})^4} & \sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{N}} \frac{\boldsymbol{t_k^3}}{(\boldsymbol{p_1} + \boldsymbol{p_2} \cdot \boldsymbol{t_k})^4} \\ \sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{N}} \frac{\boldsymbol{t_k^3}}{(\boldsymbol{p_1} + \boldsymbol{p_2} \cdot \boldsymbol{t_k})^4} & \sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{N}} \frac{\boldsymbol{t_k^4}}{(\boldsymbol{p_1} + \boldsymbol{p_2} \cdot \boldsymbol{t_k})^4} \end{pmatrix} \tag{7}$$

To determine the parameter estimation errors, according to the Cramer–Rao lower bound (Eq. (1)), the estimation error variance covariance matrix is required which is the inverse of the Fisher matrix:

$$\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)^{-1} = \frac{\mathbf{1}}{\boldsymbol{det}\left(\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)\right)} \cdot \boldsymbol{adj}\left(\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)\right) \tag{8}$$

where the adjugate matrix and determinant of the Fisher matrix are:

$$\boldsymbol{adj}\left(\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)\right) = \frac{\mathbf{1}}{\boldsymbol{\sigma^2}} \cdot \begin{pmatrix} \sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{N}} \frac{\boldsymbol{t_k^4}}{(\boldsymbol{p_1} + \boldsymbol{p_2} \cdot \boldsymbol{t_k})^4} & -\sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{N}} \frac{\boldsymbol{t_k^3}}{(\boldsymbol{p_1} + \boldsymbol{p_2} \cdot \boldsymbol{t_k})^4} \\ -\sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{N}} \frac{\boldsymbol{t_k^3}}{(\boldsymbol{p_1} + \boldsymbol{p_2} \cdot \boldsymbol{t_k})^4} & \sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{N}} \frac{\boldsymbol{t_k^2}}{(\boldsymbol{p_1} + \boldsymbol{p_2} \cdot \boldsymbol{t_k})^4} \end{pmatrix} \tag{9}$$

$$\det\left(\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)\right) = \frac{\mathbf{1}}{\boldsymbol{\sigma}^{\mathbf{4}}} \cdot \left(\sum_{k=1}^{N} \frac{\boldsymbol{t}_k^2}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \cdot \sum_{k=1}^{N} \frac{\boldsymbol{t}_k^4}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} - \left(\sum_{k=1}^{N} \frac{\boldsymbol{t}_k^3}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \right)^2 \right) \tag{10}$$

The parameter estimation errors are the square roots of the diagonal elements of the estimation error variance covariance matrix. Therefore the parameter estimation errors $\Delta \boldsymbol{p_1}$ and $\Delta \boldsymbol{p_2}$ for the Peleg model can be directly determined from $\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)^{-1}$ which gives:

$$\Delta \boldsymbol{p}_1 \geq \boldsymbol{\sigma} \cdot \sqrt{\frac{\sum_{k=1}^{N} \frac{\boldsymbol{t}_k^4}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4}}{\sum_{k=1}^{N} \frac{\boldsymbol{t}_k^2}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \cdot \sum_{k=1}^{N} \frac{\boldsymbol{t}_k^4}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} - \left(\sum_{k=1}^{N} \frac{\boldsymbol{t}_k^3}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \right)^2}} \tag{11}$$

$$\Delta \boldsymbol{p}_2 \geq \boldsymbol{\sigma} \cdot \sqrt{\frac{\sum_{k=1}^{N} \frac{\boldsymbol{t}_k^4}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4}}{\sum \boldsymbol{k}=\mathbf{1}\boldsymbol{N} \frac{\boldsymbol{t}_k^2}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \cdot \sum \boldsymbol{k} = \mathbf{1}\boldsymbol{N} \frac{\boldsymbol{t}_k^4}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} - \left(\sum \boldsymbol{k} = \mathbf{1}\boldsymbol{N} \frac{\boldsymbol{t}_k^3}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \right)^2}} \tag{12}$$

2.3. Optimal experimental design

There are different approaches in experimental design. One is equidistant design where measurements are carried out at equidistant measurement points (for example every n minutes), which is a solid approach if no information about the experiment is available. A slightly different approach is to measure more frequently at the start of the experiment and less frequently at the end, when the observed state variable shows asymptotic behaviour. Another possibility is the optimal experimental design. The general objective of the optimal design is to optimize a given objective function by variation of the experimental measurement points $\boldsymbol{x_k}$.

In this case, the goal is to find the optimal measuring points in time with respect to the analysis of the Fisher information matrix or the parameter error estimation variance covariance matrix.

The general approach to optimal experimental design is to start with the necessary rough parameter estimation values and apply the optimal design procedure. A measurement with the supposedly optimal measurement points has to be carried out. If the newly acquired parameters differ too much from the initial rough estimates, the entire process has to be repeated in sequential order with the newly calculated parameters until either the estimated and calculated parameter values match or the parameter estimation errors are considered to be sufficiently small.

2.3.1. Optimality criterion

Different optimality criteria have been developed [16] for the determination of optimal experimental design. By using the so called A-optimality criterion, the optimization goal is to minimize the trace of the parameter estimation error variance covariance matrix. In case of the Peleg model the trace is:

$$tr\left(\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)^{-1}\right) = \frac{\boldsymbol{\sigma}^2 \cdot \left(\sum \boldsymbol{k} = \mathbf{1}\boldsymbol{N} \frac{\boldsymbol{t}_k^4}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} + \sum \boldsymbol{k} = \mathbf{1}\boldsymbol{N} \frac{\boldsymbol{t}_k^2}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \right)}{\sum \boldsymbol{k} = \mathbf{1}\boldsymbol{N} \frac{\boldsymbol{t}_k^2}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \cdot \sum \boldsymbol{k} = \mathbf{1}\boldsymbol{N} \frac{\boldsymbol{t}_k^4}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} - \left(\sum \boldsymbol{k} = \mathbf{1}\boldsymbol{N} \frac{\boldsymbol{t}_k^3}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \right)^2} \tag{13}$$

When the parameter error estimation variance covariance matrix is calculated, the adjugate of the Fisher matrix is divided by the determinant, high values for the determinant result in small parameter errors. The goal of the commonly used D-optimality criterion is then to maximize the determinant of the Fisher matrix. The determinant of the Peleg model Fisher matrix is presented in Eq. (10).

When using the E-optimality criterion, the optimization target is to maximize the smallest of the eigenvalues of the Fisher matrix. For the Peleg model, the eigenvalues of the fisher matrix are:

$$\lambda\left(\boldsymbol{I}\left(\vec{\boldsymbol{p}}\right)\right) = \frac{\mathbf{1}\,\mathbf{1}}{\boldsymbol{\sigma}^{\mathbf{2}}\mathbf{2}} \cdot \left(\sum_{k=1}^{N} \frac{\boldsymbol{t}_k^4}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} + \sum_{k=1}^{N} \frac{\boldsymbol{t}_k^2}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \right) \pm \sqrt{\left(\sum_{k=1}^{N} \frac{\boldsymbol{t}_k^4}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} - \sum_{k=1}^{N} \frac{\boldsymbol{t}_k^2}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \right)^2 + \mathbf{4} \cdot \left(\sum_{k=1}^{N} \frac{\boldsymbol{t}_k^3}{(\boldsymbol{p}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{t}_k)^4} \right)^2} \tag{14}$$

The goal in the T-optimality criterion is to maximize the trace of the Fisher information matrix. But with this criterion, the non-diagonal or co-information elements of the Fisher matrix are completely ignored. Therefore the T-optimality is problematic if used alone.

2.3.2. Optimization approach

There are multiple ways to find minimum or maximum values for the various optimization criteria. The easiest and most obvious way is a numerical optimization using a constrained optimization algorithm such as for example a genetic algorithm or a particle swarm optimization algorithm.

Another way is to compute an algebraic exact solution, if possible. But it is not possible to get this kind of solution for the Peleg model because the resulting rational equation has an order higher than three which in general cannot be solved analytically.

3. Material and methods

3.1. Constraints and variables

The goal of this work was to find the optimal measurement points or times to get the smallest parameter estimation errors and to compare these results with an experimental design with equidistant measurement points in time as well as equidistant points on a logarithmic time scale. This experimental design with equidistant measurement points is usually carried out if no a priori knowledge about the experiment is available.

Some constraints for the experiment have to be defined. In the investigation the maximum duration of the experiment is determined to be 180 min (3 h). During this time span, 10 measurements are carried out. For the experiments which used equally spaced measurements, the measurements are carried out every 18 min starting at 18 min and ending at 180 min. For the experiments which used increasing time steps in between the measurements, the 10 measurement points in time were equally spaced on a logarithmic time scale from 0.25 min to 180 min (0.25, 0.52, 1.07, 2.24, 4.65, 9.67, 20.08, 41.72, 86.65 and 180 min).

Furthermore rough estimated parameter values for the Peleg model are required. The initial water content $\boldsymbol{m_0}$ was assumed to be 0.1 gramme water per gramme substrate (g_w/g_s), for $\boldsymbol{p_1}$, an initial water uptake rate of 2 (g_w/g_s) per minute was assumed, which gives 0.5 min · g_s/g_w as value for $\boldsymbol{p_1}$, and for $\boldsymbol{p_2}$ the final water content of 20 g_w/g_s was applied, which gives a parameter value of 0.05 g_s/g_w. These parameter values are in the ranges typical for model parameters of the water absorption processes of agricultural products. These constraint values and parameters are summed up in Table 1.

For the determination of the parameter estimation errors, the measurement errors are also needed. If they are unknown, reasonable values have to be guessed. In this work, a constant measurement error of 1 g_w/g_s was assumed:

$$\boldsymbol{\sigma}^2 = \left(\mathbf{1} \frac{\boldsymbol{g_w}}{\boldsymbol{g_s}} \right)^2 = \mathbf{1} \frac{\boldsymbol{g_w^2}}{\boldsymbol{g_s^2}} \tag{15}$$

Table 1
Constraints and assumed parameter values used for optimal experimental design.

$\boldsymbol{m}_0$	0.1 g_w/g_s
$\boldsymbol{p}_1$	0.5 min g_s/g_w
$\boldsymbol{p}_2$	0.05 g_s/g_w
Duration of experiment	180 min
Number of measurements $\boldsymbol{N}$	10

3.2. Optimization and programming

The numerical evaluations as well as the programming were carried out on a common desktop PC. As the analytic/algebraic calculations got very demanding computationally (~12 h) and memory wise (>100 GB Memory usage) a workstation computer (2x Intel Xeon E5 2640, 2500 MHz, 160 GB RAM, MS-Windows Server 2012) was used for these evaluations. All programming and optimization was carried out using Matlab 2014a (ver. 8.3) as well as the following toolboxes: symbolic math toolbox (ver. 6.0), optimization toolbox (ver. 7.0) and parallel computing toolbox (ver. 6.4).

In this work the A-, D-, E- and two other optimality criteria were analysed. One will be called P-criterion, where the parameter estimation error of just the first parameter $\boldsymbol{p_1}$ was minimized. Additionally a combination of the A- and P-optimality criteria was applied by minimizing the sum of the parameter estimation errors $\Delta\boldsymbol{p_i}$ divided by the absolute estimated parameter value $|\boldsymbol{p_i}|$. This will be called P_r-optimality:

$$\min\left(\frac{\Delta\boldsymbol{p_1}}{|\boldsymbol{p_1}|}+\frac{\Delta\boldsymbol{p_2}}{|\boldsymbol{p_2}|}\right) \tag{16}$$

Because pure gradient descend optimization algorithms will have difficulties solving problems with multiple local minima, a hybrid genetic algorithm (GA) was used to solve the optimization problem. For the GA, the population count was 4,000. For each optimization, up to 150 iterations were evaluated. The final result from the GA was then refined using constrained nonlinear optimization (fmincon function in Matlab).

Table 2
Results for the optimal experimental design for the Peleg model.

Optimality criterion	Measurement points (multiplicity)		Parameter estimation errors (percentage errors)	
	t_1 [min]	t_2 [min]	$\Delta\boldsymbol{p}_1$ [$min\ g_s/g_w$]	$\Delta\boldsymbol{p}_2$ [g_s/g_w]
A	6.90 (8)	180 (2)	0.041 (8.2%)	0.0021 (4.2%)
D	9.00 (5)	180 (5)	0.049 (9.8%)	0.0013 (2.6%)
E	6.90 (8)	180 (2)	0.041 (8.2%)	0.0021 (4.2%)
P	6.91 (8)	180 (2)	0.041 (8.2%)	0.0021 (4.2%)
P_r	7.31 (7)	180 (3)	0.042 (8.4%)	0.0017 (3.4%)

4. Results and discussion

4.1. Optimal experimental design results

The squared sensitivities divided by $\boldsymbol{\sigma^2}$ (SDS) for the entire process time for the parameters $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ are presented in Fig. 1. Both SDS values are zero at the beginning, because no water absorption has taken place and a measurement would not make sense. For parameter $\boldsymbol{p_1}$ the SDS value reaches its maximum at $t = 10$ min, as $\boldsymbol{p_1}$ is dependent on the initial slope of the Peleg model when $\boldsymbol{p_1} \gg \boldsymbol{p_2 t}$. When $\boldsymbol{t}$ is increasing, the slope of the Peleg model converges to zero and therefore the SDS value of $\boldsymbol{p_1}$ approaches zero as well. For $\boldsymbol{p_2}$ the SDS value rises monotonously, when $\boldsymbol{p_1} \ll \boldsymbol{p_2 t}$ the water uptake is just determined by the parameter $\boldsymbol{p_2}$ and therefore only information regarding this parameter is obtained from measurements. The SDS values are much higher for the parameter $\boldsymbol{p_2}$ which indicates that it will be more precisely determined compared to the parameter $\boldsymbol{p_1}$.

From these SDS values, one can guess optimal regions for the measurement points: In the first region the SDS value of $\boldsymbol{p_1}$ should be high and the SDS value for $\boldsymbol{p_2}$ should be low. Therefore the optimal measurement region should be after zero minutes and before the maximum of the SDS value for $\boldsymbol{p_1}$ which is at 10 min. The exact value depends on the optimality criterion. For the second region, the optimal value is obviously $\boldsymbol{t} \to \infty$, because than the value of the Peleg model depends only on $\boldsymbol{p_2}$ (see Eq. (6)). As a result the highest possible value for the measurement time is advantageous where the SDS value of $\boldsymbol{p_2}$ is high and the one of $\boldsymbol{p_1}$ is low.

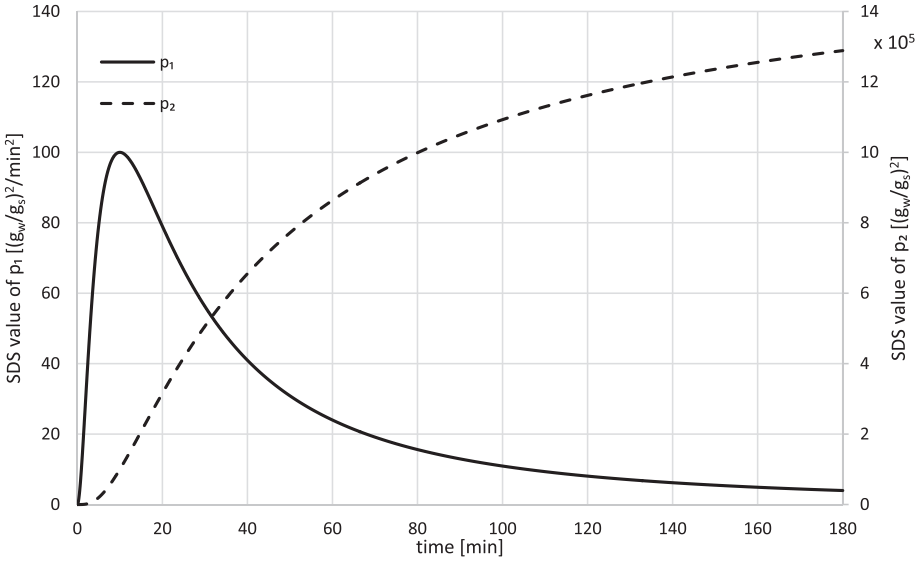

Fig. 1. Time dependence of the squared sensitivity divided by σ^2(SDS) of possible measurements of the parameters for the Peleg model.

Table 3
Parameter estimation errors, using equidistant and "logarithmic" experimental design.

Measurement points	Parameter estimation errors (percentage errors)	
t_i [min]	$\Delta \boldsymbol{p}_1$ [$min\ g_s/g_w$]	$\Delta \boldsymbol{p}_2$ [g_s/g_w]
Equidistant: {18, 36, 54, …, 180}	0.108 (21.6%)	0.0016 (3.2%)
Logarithmic: {0.25, 0.51, …, 180}	0.065 (13.0%)	0.0021 (4.2%)

In Table 2 the optimal measurement points in time and the parameter estimation errors for the Peleg model are presented depending on the different optimality criteria and assumed measurement errors.

The absolute values of $\Delta \boldsymbol{p_1}$ are always more than one magnitude higher than the $\Delta \boldsymbol{p_2}$ values. In all optimal designs just two different measurement points are determined, however with different multiplicity. The first optimal measurement point $\boldsymbol{t_1}$ is between 3.59 and 8.99 min depending on the criterion and the measurement error. The second measurement point $\boldsymbol{t_2}$ is always at the highest possible measurement time determined by the constraint ($\boldsymbol{t_2} = 180$ min). The A- and E-criterion suggest an eightfold measurement at $\boldsymbol{t_1}$ and a twofold measurement at $\boldsymbol{t_2}$. For the D-criterion the distribution of multiplicities for both measurement points is equal. For P- and P_t-criterion that is not the case.

In comparison, Table 3 shows the calculated parameter estimation errors when using the two typical, more intuitive but non optimal experimental designs such as ordinary equidistant measurement points in time as well as equidistant measurement points on a logarithmic time scale. The ordinary equidistant measurements result in a bigger error for $\boldsymbol{p_1}$ but in a smaller error for $\boldsymbol{p_2}$.

Comparing the data from Tables 2 and 3, it can be seen, that the parameter estimation error for $\boldsymbol{p_1}$ is much smaller when using an optimal design. However, the estimation error of $\boldsymbol{p_2}$ does not improve significantly. Using the A-, E- and P-optimality criteria, it even increases. At the first measurement point in the ordinary equidistant measurements ($t_1 =$ 18 min), which is much higher than any of the optimal $\boldsymbol{t_1}$ values in Table 2, $\boldsymbol{p_1}$ is smaller than $\boldsymbol{p_2} \cdot \boldsymbol{t}$ ($0.5\ \text{min} \cdot g_s/g_w < 0.9\ \text{min} \cdot g_s/g_w$) and therefore the influence of parameter $\boldsymbol{p_2}$ on the water uptake is higher resulting in a less precise parameter estimation of parameter $\boldsymbol{p_1}$ and a more precise estimation of parameter $\boldsymbol{p_2}$.

Overall using the D-optimality the smallest error for $\boldsymbol{p_2}$ is obtained while the error for $\boldsymbol{p_1}$ is only slightly worse compared to the other optimalities.

4.2. Dependency on parameter estimation precision

As the parameter values for $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ are only rough estimates, the precision of this estimation has to be evaluated. So the value for one parameter was changed, while the other was kept constant at its estimated value. This however assumes that the real values for the parameters are close to the estimated ones. If the actual value for a parameter is way off the estimation, equally spaced measurements over time might

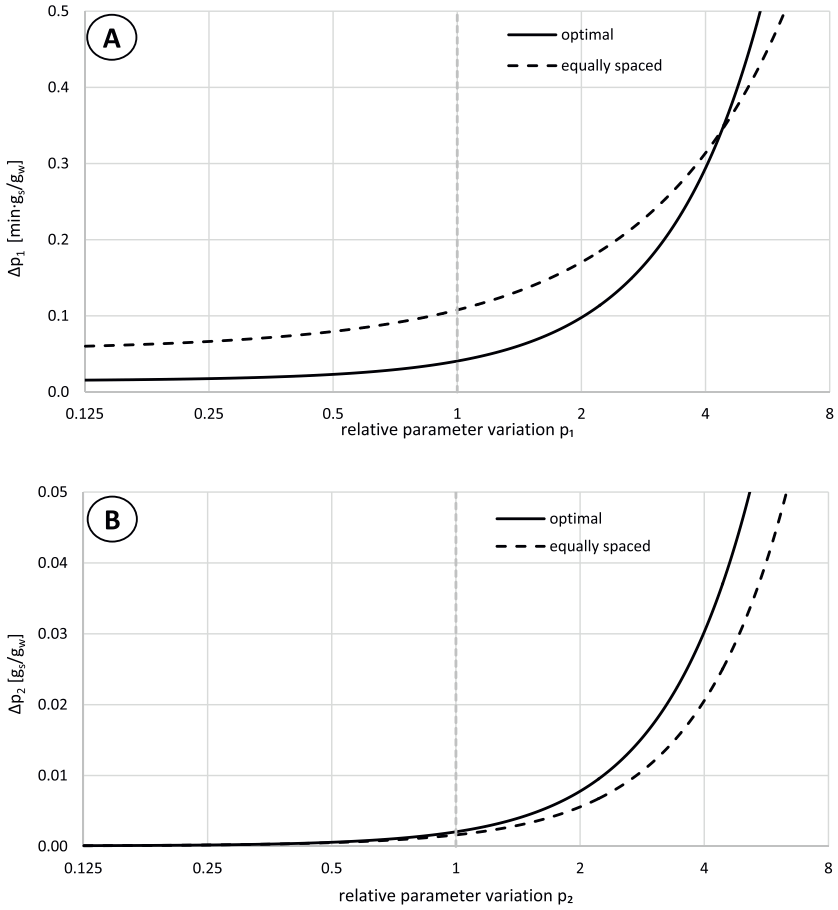

Fig. 2. Comparison of optimal to equally spaced measurement points for t as a function of the parameter values $\boldsymbol{p}_1$(A) and $\boldsymbol{p}_2$(B). While one parameter was evaluated, the other parameter was fixed to its initial value. The dashed vertical lines mark the actual values for the parameters. (A-optimality,$\sigma^2_{\hat{\mathbf{t}}_1}$).

Table 4
Relative parameter variation range in which optimal experimental design is superior to equidistant experimental design.

Optimality criterion	Relative variation of p_1		Relative variation of p_2	
	Lower	Upper	Lower	Upper
A	<0	4.38	0.19	0.31
D	<0	3.99	0.14	>10
E	<0	4.38	0.19	0.31
P	<0	4.39	0.19	0.31
P_r	<0	4.22	0.15	0.80

provide better result/smaller parameter estimation errors than the "optimal" measurement points. Fig. 2 shows the result of this evaluation for both parameter estimation errors using the A-criterion.

It can be seen, that forΔp_1, the optimal design gets worse than the equidistant experimental design if the value for p_1 is about 4.4 times bigger than its rough estimation. For Δp_2 the optimal design seems to be worse than the equidistant experimental design. This is in accordance to the results presented in Table 2.

In Table 4 the results for this comparison of optimal and equidistant experimental design are presented for all evaluated optimality criteria and measurement error types.

It is obvious, that for the A-, E- and P-optimality criteria the optimal experimental design is only useful to reduce the first parameter estimation error Δp_1. With the D- and P_r-criterion however, which used relative parameter errors, the optimal experimental design can also be beneficial to decrease the second parameter estimation error Δp_2 as well, depending on the estimation precision of p_2 and the assumed measurement error. But when the D- and P_r-criteria are used, the reduction of Δp_1 is considerably worse than when using A-, E- or P-optimality.

4.3. Approximation for optimal experimental design using an algebraic solution

As one can see from Table 2, the second optimal measurement point in time t_2 is always equal to the upper constraint of the optimization space, which was 180 min in this investigation. If a function F can be found to relate the t_1 value to the t_2 value, the optimal experimental design process is significantly simplified as no optimization is necessary and the optimal experimental design result is obtained immediately.

To find an algebraic solution for the optimal design using the A-optimality it is necessary to calculate the first and second partial derivative of the optimality criterion for the measurement point t_1 or t_2. Eq. (17) shows the conditions to solve to find a minimum in the optimality criterion A:

$$\frac{\partial tr\left(I\left(\vec{p}\right)^{-1}\right)}{\partial t_1 = 0} \cup \frac{\partial^2 tr\left(I\left(\vec{p}\right)^{-1}\right)}{\partial (t_1)^2} \geq 0 \tag{17}$$

By analysing Eq. (13) it appears, that the first derivative cannot be solved for t_1 as it is a rational function of sixth order. However, it is possible to get the roots of this derivative by numerical approximation using, for example, the Newton method. Fig. 3 shows the results of the repeated numerical calculation of the roots for t_1 while varying t_2.

As the first derivative has only one root between t = 0 and t = t_2, there is no necessity to check if the second derivative is higher than zero. The only root must lead to a minimum.

The point is, that by this approach one can generate the aforementioned approximation function F for t_1, which depends on t_2, p_1, p_2 and on an additional parameter a.

$$t_1 \approx F(t_2, p_1, p_2, a) \tag{18}$$

The additional parameter a depends on the optimality criterion, on the multiplicity of the two optimal measuring points t_1 and t_2 and on the assumed measurement error. The function, which can approximately describe the dependency of t_1 and t_2, is coincidentally quite similar to the Peleg model function, but with only one parameter (zero parameters for D-optimality):

$$t_1 = \frac{t_2}{\left(a + 1 + a \cdot \frac{p_2}{p_1} \cdot t_2\right)} \tag{19}$$

The value of a depends on the optimality criterion. If the criterion implies no weighting on the parameter errors as the D-optimality, a becomes 1 and this approximation is in fact an exact solution. It is obtained analytically by calculating the roots of the derivative of $det\left(I\left(\vec{p}\right)\right)$ which is a rational function of third order (the roots are zero, t_2 and the Eq. (19) with $a = 1$). For A-, E- and P-optimality a becomes ~1.31 and for the P_r-optimality a is ~1.24

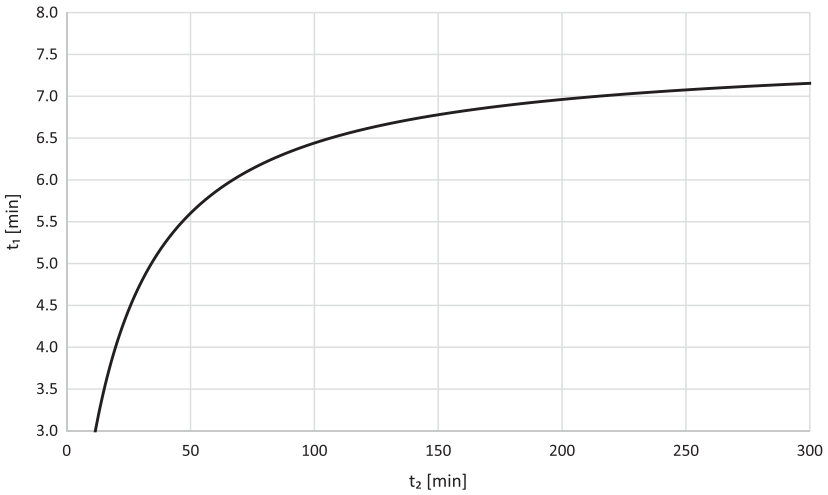

Fig. 3. Optimal value of t_1 as function of t_2. Values were calculated by numerical calculation using Eq. (18) (A-optimality).

Table 5
Results for the optimal experimental design for the Peleg model using the approximation function. The differences to the values acquired by the GA are shown in the brackets.

Optimality criterion	t_1 [min]	$\Delta \boldsymbol{p}_1$ [$min\ g_s/g_w$]	$\Delta \boldsymbol{p}_2$ [g_s/g_w]
A	6.93 (+0.03)	0.041 (+2e−5)	0.0021 (−6e−6)
D	9.00	0.049	0.0013
E	6.93 (+0.03)	0.041 (+2e−5)	0.0021 (−6e−6)
P	6.93 (+0.02)	0.041 (+3e−5)	0.0021 (−4e−6)
P_r	7.32 (+0.01)	0.042 (+3e−6)	0.0024 (−8e−7)

As $\boldsymbol{\sigma^2}$ is factored out and is not in Eq. (19), it will not affect the result of the optimal experimental design. When using the approximation to calculate $\boldsymbol{t_1}$, the values and the differences to the values gathered by the normal optimal experimental design using a GA are shown in Table 5. It can be seen, that the differences for the approximation function are mostly negligible as they are very low.

5. Conclusion

In this work the optimal design of experiments for the Peleg model was investigated. For this the parameter estimation errors have been calculated using the Cramer–Rao lower bound. Various optimality criteria and changing precision of roughly estimated parameters were analysed and compared to equidistant measurements on ordinary and logarithmic time scale.

For the ordinary and logarithmic equidistant experimental design, the relative parameter estimation errors $\Delta \boldsymbol{p_1}$ and $\Delta \boldsymbol{p_2}$ were very different with 22% and 13% of the parameter value for $\boldsymbol{p_1}$, and only 3% and 4% of the value of $\boldsymbol{p_2}$. So it is obvious that in the minimization of the parameter estimation errors, $\Delta \boldsymbol{p_1}$ leaves much more room for optimization. For that reasons it is not surprising that, when comparing the optimal design results to the ordinary equidistant results, especially the estimation error $\Delta \boldsymbol{p_1}$ of the Peleg model could be improved significantly by up to 62% (A-,E- and P-optimality). For the second parameter, the estimation error could not be reduced that much. The best reduction was 19% (D-optimality). In most cases the estimation error $\Delta \boldsymbol{p_2}$ was even slightly bigger with the optimal design. This is also as expected as the optimizer "sacrifices" a little increase in $\Delta \boldsymbol{p_2}$ to drastically decrease $\Delta \boldsymbol{p_1}$.

Another way to explain this behaviour is that the problem itself is a multi-objective problem where multiple parameter estimation errors should be minimized at the same time. This gets apparent when looking at Tables 2 and 3. All but the D-optimality criterion result in objective functions which imply different weightings to these parameter estimation errors and therefore give different results for the optimal design and "optimal" parameter estimation. For the A- and E-optimality, the weighting happens as the absolute parameter variances are added together. If one value is significantly higher than the others, the others will become irrelevant even though the relative errors may all be more or less the same. One simple solution to this problem is to use the P_r-optimality, where the parameter estimation errors get divided by their parameter value. So the weighting is applied due to the relative parameter estimation errors. If a completely unbiased and unweighted optimality is desired, the D-optimality should be used. If the goal of the optimal experimental design is to precisely determine just one parameter, the P-optimality simplifies the problem to a single objective function and is preferable.

Also a simple approximation method for the optimal experimental design for the Peleg model was developed in this work. Therefore no optimization is required for the Peleg model anymore. By applying the proposed function (Eq. (19)) the optimal experimental design result is obtained immediately. This might be useful when repeated water absorption measurements with different materials are planned. Until now classical equidistant experimental design was the usual solution as it was far too time-consuming to calculate an optimal experimental design for each compound. When using the common D-optimality, the approximation and the optimal design using for example a genetic algorithm give the exact same results, but the approximation can be calculated much faster literally using pen and paper. The differences between the normal optimal experimental design approach using an optimization algorithm and the introduced approximation function are mostly negligible, although there is still room for improvements.

Higher values for $\boldsymbol{t_2}$ will generally result in a more precise parameter estimation for the Peleg model. But as $\boldsymbol{t_2}$ is equal to the length of the entire experiment, a compromise has to be made between parameter estimation precision and experiment duration. This however is not a problem of the approximation but the Peleg model itself.

Conflict of interest

The authors declare that there are no conflicts of interest.

References

[1] C. Gaiani, P. Schuck, J. Scher, S. Desobry, S. Banon, Dairy powder rehydration: influence of protein state, incorporation mode, and agglomeration, J. Dairy Sci. 90 (2007) 570–581.
[2] G.J.O. Martin, R.P.W. Williams, D.E. Dunstan, Comparison of casein micelles in raw and reconstituted skim milk, J. Dairy Sci. 90 (2007) 4543–4551.
[3] J. Butcher, N.L. Stenvert, Conditioning studies on Australian wheat. I. The effect of conditioning on milling behaviour, J. Sci. Food Agric. 24 (1973) 1055–1066.
[4] F.D. Montanuci, C.A. Perussello, Luiz Mario de Matos Jorge, Regina Maria Matos Jorge, Experimental analysis and finite element simulation of the hydration process of barley grains, J. Food Eng. 131 (2014) 44–49.
[5] P.A. Sopade, E.S. Ajisegiri, M.H. Badau, The use of Peleg's equation to model water absorption in some cereal grains during soaking, J. Food Eng. 15 (1992) 269–283.
[6] A.R. Piergiovanni, Kinetic of water adsorption in common bean: considerations on the suitability of Peleg's model for describing bean hydration, J. Food Process. Preserv. 35 (2011) 447–452.
[7] M. Turhan, S. Sayar, S. Gunasekaran, Application of Peleg model to study water absorption in chickpea during soaking, J. Food Eng. 53 (2002) 153–159.
[8] K.D. Dolan, D.K. Mishra, Parameter estimation in food science, Annu. Rev. Food Sci. Technol. 4 (2013) 401–422.
[9] A. Mannarswamy, S.H. Munson-McGee, R. Steiner, P.K. Andersen, D-optimal experimental designs for Freundlich and Langmuir adsorption isotherms, Chemom. Intell. Lab. Syst. 97 (2009) 146–151.
[10] F. Ataíde, B. Hitzmann, When is optimal experimental design advantageous for the analysis of Michaelis–Menten kinetics? Chemom. Intell. Lab. Syst. 99 (2009) 9–18.
[11] Brecht M.R. Donckels, Dirk J.W. De Pauw, P.A. Vanrolleghem, B. de Baets, Performance assessment of the anticipatory approach to optimal experimental design for model discrimination, Chemom. Intell. Lab. Syst. 110 (2012) 20–31.
[12] G. Franceschini, S. Macchietto, Model-based design of experiments for parameter precision: state of the art, Model Based Exp. Anal. 63 (2008) 4846–4872.
[13] A. Boukouvalasa, D. Cornforda, M. Stehlíkb, Optimal design for correlated processes with input-dependent noise, Comput. Stat. Data Anal. 71 (2014) 1088–1102.
[14] J. Songok, P. Salminen, M. Toivakka, Temperature effects on dynamic water absorption into paper, J. Colloid Interface Sci. 418 (2014) 373–377.
[15] M. Peleg, An empirical model for the description of moisture sorption curves, J. Food Sci. 53 (1988) 1216–1217.
[16] F. Pukelsheim, Optimal Design of Experiments, Wiley, New York NY u.a, 1993..

3.2 Optimal design of experiments and measurements of the water sorption process of wheat grains using a modified Peleg model

By Olivier Paquet-Durand, Viktoria Zettel, Reinhard Kohlus and Bernd Hitzmann.

Published 2015 in **Journal of Food Engineering**.

Volume 165, Pages 166 - 171

Journal of Food Engineering 165 (2015) 166–171

Contents lists available at ScienceDirect

Journal of Food Engineering

journal homepage: www.elsevier.com/locate/jfoodeng

Optimal design of experiments and measurements of the water sorption process of wheat grains using a modified Peleg model

O. Paquet-Durand *, V. Zettel, R. Kohlus, B. Hitzmann

Department of Process Analytics and Cereal Science, Institute of Food Science and Biotechnology, University of Hohenheim, Garbenstr. 23, 70599 Stuttgart, Germany

ARTICLE INFO

Article history:
Received 5 May 2015
Received in revised form 12 June 2015
Accepted 17 June 2015
Available online 18 June 2015

Keywords:
Optimal design of experiment
Fisher information
Cramer-Rao lower bound
Water absorption
Peleg equation
Wheat grains

ABSTRACT

Wheat grains consist of three major components, the bran layer, the endosperm and the germ, with very different water sorption kinetic. The original two parameter Peleg model cannot describe a water sorption process well for such heterogeneous compounds. Therefore, the model was modified to account for the two biggest fractions, the endosperm and the bran layer. This modified model has four parameters and can be used to accurately describe the hydration process of wheat grains.

Two experiments were carried out, an initial experiment to get rough parameter values and a second experiment, which was designed optimally by using the Cramer-Rao lower bond method. The percentage parameter errors for the four parameters of the modified model were reduced from 669%, 24%, 12%, and 2.4% to 38%, 5.4%, 4.5% and 1.9% respectively. The presented results demonstrate the advantage of optimal design of experiments.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Hydration kinetics in food technology are key physicochemical properties. The knowledge of moisture content and moisturizing kinetics is vitally important to determine the chemical and physical properties of food products and their shelf life.

In this work, the water absorption kinetics of wheat grains were investigated and an optimal experimental design was applied. The actual water content of wheat grains due to storage is too low for optimal milling. During the milling process the bran and the germ have to be separated efficiently from the endosperm, because the economic values are related to their purity. The water addition exaggerates the difference of the different parts of the wheat kernel and simplifies the separation process. The hydration kinetics for the endosperm is rather slow compared to the germ and the bran (Delcour and Hoseney, 2010) and a hydrated bran is very flexible and will stay mostly intact during milling and can then be sieved out. Therefore, a hydration step, the tempering, is applied routinely before milling to produce optimal quality. A more detailed knowledge about the hydration kinetics of the components of the wheat grains is therefore required to prepare an optimal separation process.

There are fundamental approaches to water absorption kinetics. Cunningham et al. (2007) and Munson-McGee et al. (2011a,b) used models for water uptake based on Fick's diffusion laws. However, models based on the diffusion law are usually very complex and not very convenient for computing in most situations. Therefore, Peleg (1988) suggested an empiric two parameter model to describe water absorption curves. For example Sopade et al. (1992) and Maskan (2001) used this model to accurately describe the water absorption of cereal grains and wheat respectively. It is still very common due to its simplicity and ideal for an optimal design of experiments due to its modest computational requirements.

Optimal experimental design is usually carried out to get the maximal amount of information from an experiment with the least amount of effort. This goal is achieved by variation of measurement times and measurement locations or other process variables such temperature or pH. One specific and also quite common goal of an optimal experimental design can be to minimize parameter estimation errors. Ataíde and Hitzmann (2009) used an optimal experimental design to analyze enzyme kinetics in a stirred tank reactor. They proved that fed-batch processes will always result in smaller parameter estimation errors than for normal batch processes. Franceschini and Macchietto (2008) presented a list of various recent model based optimal experimental design applications from chemical kinetics to biological modelling. Sánchez et al. (2012) presented a method for a pareto-optimal design. They showed that it is possible to get a non-optimal but still good experimental design regarding multiple criteria at the same time.

* Corresponding author.
E-mail address: o.paquet-durand@uni-hohenheim.de (O. Paquet-Durand).

http://dx.doi.org/10.1016/j.jfoodeng.2015.06.025
0260-8774/© 2015 Elsevier Ltd. All rights reserved.

According to Dolan and Mishra (2013), the importance of accurate parameter error estimation is often underestimated in food science. In this work, the optimal design of experiments to improve the parameter estimation error of a modified Peleg model was investigated. To do so, an initial experiment for water sorption of wheat grains was carried out to roughly estimate the parameters. Afterwards these initial parameters were used to carry an optimal experimental designs using the Cramer-Rao lower bound method.

2. Performing optimal design of experiments

The goal of optimal design of experiments is to determine optimal measurement points or the evolution of process variables (such as batch or fed-batch mode) for an experiment, in a way that the variances of model parameters calculated from corresponding measurements are as small as possible. Therefore, one requirement is a theoretical model describing the process under consideration.

2.1. The theoretical model for water sorption

A very common empiric model to describe water sorption processes is the two parameter empiric Peleg model (1988) which is defined as:

$$M(t,p_1,p_2,m_0) = m_0 + \frac{t}{p1 + p_2 t} \tag{1}$$

m_0 is the starting water content of the sample, $\frac{1}{p_1}$ is the initial water uptake rate and $\frac{1}{p_2} + m_0$ equals the final water content. If the relative moisture increase is measured, m_0 can be set to zero and therefore excluded.

Wheat grains consist of 3 major components with very different water absorption kinetics: The bran layer, the endosperm and the germ. The germ was ignored as its part of the total mass of a grain is about 3%. The bran layer absorbs small amounts of water very quickly and the endosperm absorbs large amounts of water in a slow manner by comparison (Delcour and Hoseney, 2010). The different amounts of water uptake are due to the different mass amounts. The ratio of bran to endosperm mass is roughly 1:10–1:5 (Erling and Botterbrodt, 2008).

The normal Peleg model works quite well for homogenous materials (Cunningham et al., 2007; Shafaei et al., 2014; Turhan et al., 2002) but in case of the heterogeneous wheat grains, it is not very accurate. So it has to be modified to account for the heterogeneity. The most obvious way to account for this is to model the water sorption of the bran layer and the endosperm individually:

$$M_{bran}(t,p_{1_b},p_{2_b},m_{0_b}) = m_{0_b} + \frac{t}{p_{1_b} + p_{2_b} t} \tag{2}$$

$$M_{endo}(t,p_{1_e},p_{2_e},m_{0_e}) = m_{0_e} + \frac{t}{p_{1_e} + p_{2_e} t} \tag{3}$$

where m_{0_b} p_{1_b} and p_{2_b} are the Peleg model parameters of the bran, and, m_{0_e} p_{1_e} and p_{2_e} are the Peleg model parameters of the endosperm. The total water absorption of a wheat grain is then the sum of the two corresponding terms:

$$M_{grain}(t,p_{1_b},p_{2_b},p_{1_e},p_{2_e}) = \frac{t}{p_{1_b} + p_{2_b} t} + \frac{t}{p_{1_e} + p_{2_e} t} \tag{4}$$

The modified Peleg model now consist of two sets of Peleg parameters, one for each compound. In this contribution the m_0 values are set to zero because just the absorbed water in relation to the initial wet grain mass is considered here.

2.2. Parameter error estimation

For the determination of the model parameters a least squares fit to measurement data is performed. To estimate parameter errors, the Cramer-Rao inequality or lower bound (CRlb) can be applied (Bos, 2007). It states that the lower bound of the variance of an estimated parameter p_i is equal or higher to the corresponding diagonal element of the inverse of the Fisher information matrix:

$$var(p_i) \geqslant [I(\vec{p})^{-1}]_{ii} \tag{5}$$

Here $var(p_i)$ is the variance of the ith parameter, $I(\vec{p})^{-1}$ is the inverse of the Fisher information matrix, and $[\ldots]_{ii}$ is the ith diagonal element of the matrix. If the measurements are described by a function $f(t_k,\vec{p})$ with the independent variable t_k and the parameters $\vec{p}$ the ij-element of the Fisher information matrix is defined as:

$$I(\vec{p})_{ij} = \frac{1}{\sigma^2} \cdot \sum_{k=1}^{N} \frac{\partial f(t_k,\vec{p})}{\partial p_i} \frac{\partial f(t_k,\vec{p})}{\partial p_j}. \tag{6}$$

σ is the measurement error (standard deviation of the measurements in this case) and N is the number of measurements. For the water absorption process the model for the measurement variable is given as:

$$f(t_k,\vec{p}) = M_{grain}(t_k,p_{1_b},p_{2_b},p_{1_e},p_{2_e}) \tag{7}$$

t is the independent variable for the measurement time. The ij-elements of the Fisher information matrix are therefore calculated as follows:

$$I(\vec{p})_{ij} = \frac{1}{\sigma^2} \cdot \sum_{k=1}^{N} \frac{\partial M_{grain}(t_k,p_{1_b},p_{2_b},p_{1_e},p_{2_e})}{\partial p_i} \cdot \frac{\partial M_{grain}(t_k,p_{1_b},p_{2_b},p_{1_e},p_{2_e})}{\partial p_j} \tag{8}$$

As one can see, the change of the M_{grain} function with respect to a change of the parameter at the measurement points t_k has an influence on the values of the element of the Fisher information matrix. If the change in M_{grain} with respect to a change of a parameter is high, the information is high, if on the other hand the function does not depend on the parameter at a measurement point the corresponding value of the information is low.

The full matrix is calculated as follows:

$$I(p) = \frac{1}{\sigma^2} \cdot \sum_{k=1}^{N} \begin{pmatrix} \frac{t_k^2}{(p_{1_b}+p_{2_b}t_k)^4} & \frac{t_k^3}{(p_{1_b}+p_{2_b}t_k)^4} & \frac{t_k^2}{(p_{1_b}+p_{2_b}t_k)^2(p_{1_e}+p_{2_e}t_k)^2} & \frac{t_k^3}{(p_{1_b}+p_{2_b}t_k)^2(p_{1_e}+p_{2_e}t_k)^2} \\ \frac{t_k^3}{(p_{1_b}+p_{2_b}t_k)^4} & \frac{t_k^4}{(p_{1_b}+p_{2_b}t_k)^4} & \frac{t_k^3}{(p_{1_b}+p_{2_b}t_k)^2(p_{1_e}+p_{2_e}t_k)^2} & \frac{t_k^4}{(p_{1_b}+p_{2_b}t_k)^2(p_{1_e}+p_{2_e}t_k)^2} \\ \frac{t_k^2}{(p_{1_b}+p_{2_b}t_k)^2(p_{1_e}+p_{2_e}t_k)^2} & \frac{t_k^3}{(p_{1_b}+p_{2_b}t_k)^2(p_{1_e}+p_{2_e}t_k)^2} & \frac{t_k^2}{(p_{1_e}+p_{2_e}t_k)^4} & \frac{t_k^3}{(p_{1_e}+p_{2_e}t_k)^4} \\ \frac{t_k^3}{(p_{1_b}+p_{2_b}t_k)^2(p_{1_e}+p_{2_e}t_k)^2} & \frac{t_k^2}{(p_{1_b}+p_{2_b}t_k)^2(p_{1_e}+p_{2_e}t_k)^2} & \frac{t_k^3}{(p_{1_e}+p_{2_e}t_k)^4} & \frac{t_k^4}{(p_{1_e}+p_{2_e}t_k)^4} \end{pmatrix} \tag{9}$$

The diagonal elements of the Fisher information matrix are called sensitivities, carrying the information with respect to each corresponding parameter.

2.3. Optimal design of experiments

If no a priori information about an experiment or its expected results are available, the usual approach is to perform an initial experiment were the measurements are taken equidistantly or in any other reasonable fashion. With these measurement values, the model parameters are determined. However, due to possible suboptimal measurement points the parameter accuracy might not be high. Using these mostly non-accurate parameter values, the optimal experimental design can be calculated using an optimality criterion, which is discussed below. The parameter estimation errors are minimized by changing the measurement points. Carrying out the new experiment, improved parameter values can be calculated. If the new estimated parameters deviate by a high amount from the first estimation, the optimal design procedure should be repeated until the parameter errors are suitable or the change in the design is negligible. A cost-effectiveness consideration should be made after each iteration as the cost or effort increases with each iteration linearly but the benefit becomes smaller and smaller. In most cases, if the initial parameter guess is reasonable, one single iteration will give sufficient results with the least amount of effort.

When applying the optimal experimental design using the Fisher information and the CRlb, there are many ways to determine the quality of the parameter estimation (Pukelsheim, 1993). The most commonly used optimality criterion is probably the D-optimality. In a D-optimal design of experiment the goal is to maximize the determinant of the Fisher information matrix. If the variance in the relative parameter estimation error is high as in this contribution it is usually advantageous to minimize the sum of relative parameter errors which is called P_r criterion (Paquet-Durand et al., 2014).

$$min\left(\frac{\Delta p_{1_{bran}}}{p_{1_{bran}}}+\frac{\Delta p_{2_{bran}}}{p_{2_{bran}}}+\frac{\Delta p_{1_{endo}}}{p_{1_{endo}}}+\frac{\Delta p_{2_{endo}}}{p_{2_{endo}}}\right) \tag{10}$$

For the CRlb method this is the sum of the diagonal elements of the inverse Fisher information matrix divided each by the corresponding parameter value.

3. Material and methods

3.1. Experiments

Some constraints for the optimal design of experiments had to be specified. Here the maximum duration of the experiments was set to 48 h (2 days) and the number of measurements to be carried out specified to be 12.

Altogether two experiments were performed. First, the initial non-optimal experiment for water absorption whose measurement values are used for the rough estimation of the parameter values of the modified Peleg model. The 12 measurements (in triplicate) were taken at the following times: 0.5 h, 1 h, 2 h, 4 h, 6 h, 8 h, 12 h, 16 h, 20 h, 24 h, 36 h and 48 h. The optimized second experiment was carried out according to the calculation of optimal time points by the optimal design procedure.

For a measurement 5–10 g wheat grains where deposited in demineralized water at 30 °C. After the storage time the grains were dried with paper towels and weighted again. To calculate the relative water uptake the difference of the final and initial wet mass was divided by the initial wet mass.

3.2. Programming and optimization procedure

For the programming and optimization Matlab 2014a (ver. 8.3) as well as the following toolboxes were used: symbolic math toolbox (ver. 6.0) and optimization toolbox (ver. 7.0).

For the optimization the genetic algorithm implementation (GA) of Matlab was used. The population count of the GA was 150 and up to 150 iterations were evaluated. The mutation was specified to be like Gaussian noise with 100% standard deviation in the first iteration, linearly decreasing to 0% standard deviation in the last iteration. The crossover fraction was set to 0.8.

3.3. Measurement error

The measurement error has to be known to calculate the parameter errors with the CRlb method. If the exact measurement errors are unknown, a reasonable value has to be estimated instead. In this contribution, the error of the used scale was given as ±0.001 g. The error in the measurement time is estimated to be 5 s. A significant but unknown error is introduced by the paper towel drying of the wheat grains. Form the known error sources a valid and accurate estimation of the measurement error was not feasible. Instead a constant measurement error of 0.78% was assumed, which is the average standard deviation of the performed measurements:

$$\sigma^2 = (0.78\%)^2 = 0.61\%^2 \tag{11}$$

4. Results & discussion

In Fig. 1 the results of the initial non-optimal water uptake measurements as well as the fitted original 2 parameter Peleg model and the modified 4 parameter Peleg model are shown. All water uptake values are relative to the initial wet weight of the grains. Also shown are the two individual water uptake progressions of the bran layer and the endosperm of the modified model, which correspond to the two major compounds of the grain. The root mean square error of the modified Peleg model is 0.45%. The root mean square error of original two parameter Peleg model is 2.51%. It can also be seen, that the original Peleg model involves a systematic error.

It can be seen, that the outer bran layer absorbs the water very fast and is saturated after a few minutes whereas the endosperm absorbs the water much slower, but the total amount of absorbed water is much higher. From the fitted model one can see that even

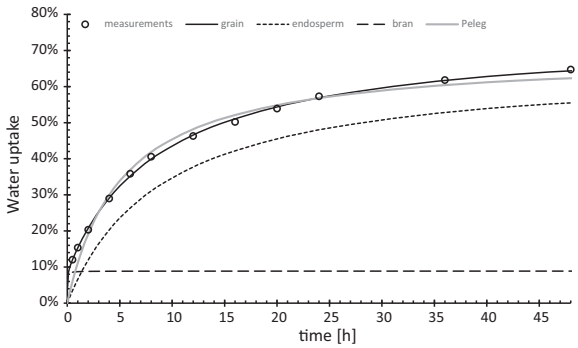

Fig. 1. Relative water uptake of the wheat grains at 30 °C for the initial measurements and the fitted modified 4 parameter Peleg model for the entire grain as well as the bran and endosperm partial water uptakes. For comparison the classical 2 parameter Peleg model fit is shown as well, indicated as "Peleg".

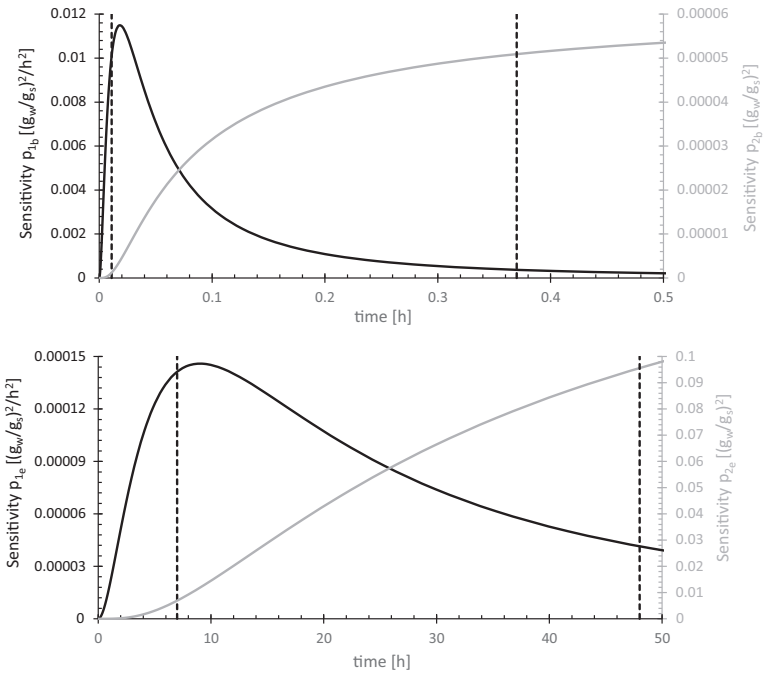

Fig. 2. Sensitivities of the four parameters of the modified Peleg model. The timescale for the sensitivities of the bran parameters is much smaller as the water absorption process in the bran happens much faster. The vertical dotted lines indicate the four optimal measurement points.

after 48 h the endosperm is still not completely saturated with water.

Using the parameter values calculated by fitting the modified Peleg model to the initial measurements, the optimal experimental design was carried out. In Fig. 2 the sensitivities of the four parameters of the modified Peleg model are shown. The dotted vertical lines indicate optimal measurement points obtained from the optimization procedure using the aforementioned genetic algorithm. It can be seen, that the optimal measurement points in time are not exactly at the maximum values of the corresponding sensitivity. The reason is that for an ideal determination of a parameter not only the sensitivity of the actual parameter should be as high as possible but also the sensitivities of all other parameter should be as low as possible at the same time. The ideal measurement points for p_{1_b} and p_{1_e} would be expected to be just slightly lower than the position of their sensitivity maxima. The optimal measurement point for p_{2_e} is 48 h which is the specified maximum duration of the experiment. The ideal measurement point would be $+\infty$ because the sensitivities for p_{1_b} and p_{1_e} will become smaller and smaller whereas the sensitivity of p_{2_e} will still increase faster than p_{2_b}. The ideal measurement point for p_{2_b} would expected to be between the sensitivity maxima of p_{1_b} and p_{1_e}.

The optimal measurement points are also shown in Table 1 as well as their measured relative water uptake values. For the sake of completeness the measurement points and water uptake values of the initial experiment are shown too.

The modified Peleg model has 4 parameters, so in the optimal case there must be 4 measurement points. The ideal measurement points were 0.011 h, 0.37 h, 7 h and 48 h when using the P_r criterion.

In Fig. 3 the resulting optimal measurements as well as the new fitted modified Peleg model are presented. The scale of the time axis has been reduced to improve the visibility of the first few hours of the experiment.

It can be seen that the very first measurement (0.011 h) was conducted very early on in the water absorption process where the bran has absorbed half of its maximum water capacity. The second optimal measurement time (0.37 h) is when the water

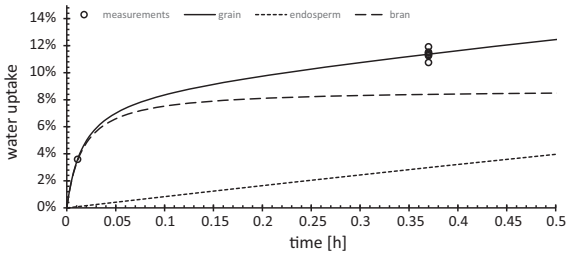

Fig. 3. Water uptake for the optimal measurements and the fitted modified Peleg model as well as the partial uptake of the bran and endosperm.

Table 1
Resulting measurement time points of the two experiments as well as the corresponding measured water uptake values.

Initial experiment	Time [h]	0.5	1	2	4	6	8	12	16	20	24	36	48
	Water uptake [%]	12.0	15.4	20.3	29.0	35.8	40.6	46.3	50.2	53.9	57.3	61.8	64.7
Optimally designed experiment	Time [h]	0.011	0.37	0.37	0.37	0.37	0.37	0.37	7	7	7	7	48
	Water uptake [%]	3.6	11.3	11.5	10.8	11.5	11.9	11.3	39.9	39.7	39.9	39.3	64.7

Table 2
Parameter values and their estimation errors for the two performed experiments.

Parameter	Initial experiment		Optimally designed experiment	
	Value	Error	Value	Error
p_{1_b} [h g_w/g_s]	0.207	±1.382 (668.5%)	0.189	±0.072 (38.3%)
p_{2_b} [g_w/g_s]	11.279	±2.667 (23.6%)	11.404	±0.619 (5.4%)
p_{1_e} [h g_w/g_s]	13.669	±1.569 (11.5%)	11.841	±0.534 (4.5%)
p_{2_e} [g_w/g_s]	1.514	±0.036 (2.4%)	1.541	±0.030 (1.9%)

absorption process for the bran is almost finished. At the third optimal measurement time (7 h), the endosperm has absorbed about 50% of its maximal water capacity. The last measurement point would be, when the endosperm reached it terminal water capacity. But this will not be the case in the time span considered for the experiment so the closest allowed value was selected which was the upper limit of 48 h.

Table 2 shows the parameter values as well as the corresponding estimation errors for the non-optimal experiment as well as for the optimal experiment. Both corresponding values do not deviate much from each other however, their errors do.

The parameter estimation errors for the initial experiment are much higher than for the optimally designed experiment, which was to be expected. The error for p_{1_b} is more than six times bigger than the value of p_{1_b} itself. For the other parameters, the errors are not quite as big. The reason why the error of p_{1_b} is as big is due to the fact that the water absorption of the bran happened very fast within the first minutes. The first actual measurement in the non-optimal experiment was taken after 30 min, which was too late for a good determination of the p_{1_b} parameter. The results of the optimal design indicated that the ideal measurement time for the determination of p_{1_b} is at about 0.011 h (36 s). The optimally designed experiment therefore resulted in a 17.4 times smaller relative parameter estimation error for p_{1_b}. The relative estimation errors for the other three parameters are also significantly lower by a factor of 4.4 for p_{2_b}, 2.5 for p_{1_e} and 1.3 for p_{2_e}.

As one can see, the parameter p_{1_b} is much smaller as p_{1_e} again indicating the faster water absorption by the bran. If the values of p_{2_b} and p_{2_e} are compared, which determine the overall water absorption of each compound, the smaller value of p_{2_e} by a factor of 7.4 indicate the higher amount of water absorption of the endosperm compared to the bran layer. This is well in the region of 1:10–1:5 for the bran-endosperm mass ratio, although the specific water uptake ability of the two major components might be slightly different.

Another optimal design of experiments step was applied with the newly calculated parameters of the modified Peleg model. The ideal measurement points of this second iteration are 1×0.011 h, 6×0.35 h, 4×6.3 h and 1×48 h. The only significant difference is in the third measurement point, which went from 7 h to 6.3 h. However, a simulation showed that a third experiment at these points would probably not decrease the estimation errors of any parameter by more than 2%. Therefore, a third experiment was not considered worth the effort.

For both models the model error at a certain point in time can be calculated by error propagation using the calculated parameter errors. It is obvious, that smaller parameter errors will result in a smaller model error (data not shown). However, from an engineering point of view, it might be more interesting how the error of the time will evolve with the time necessary to reach a certain moisture level of the grain? In Fig. 4 these errors are presented for the initial model as well as the optimal model.

It can be seen, that the error in time to reach a defined moisture is improved significantly. The error of the initial model with the rather large parameter errors is significantly higher (about three to ten times) as the one with the optimized model.

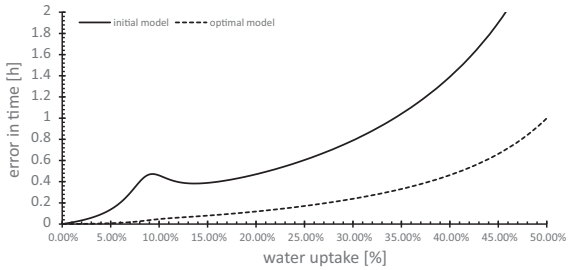

Fig. 4. Error in time prediction as a function of the water uptake for the initial and the optimized modified Peleg model.

5. Conclusion

For an overall process optimization, theoretical models as well as rough values for the model parameters must be known. In this investigation, the water absorption process of wheat grains was considered. At first an experiment with 12 moisture measurements in triplicate was performed and the measurement times were chosen based on experience. It could be shown that the modified Peleg model is very suitable to describe the measured moisturizing process of fresh wheat grains, where the normal Peleg equation left much to be desired as the model error could be reduced from 2.51% in case of the normal Peleg model to 0.45% in case of the modified Peleg model.

Afterwards an optimal design of experiments was carried out to determine the four model parameter as accurate as possible. A second experiment again with 12 measurements in triplicate was carried out where the optimal measurement points in time were calculated based on the Cramer-Rao lower bond. The parameter estimation errors for the model equation could be reduced by a factor of up to 17.4 in case of the $p_{1_{bran}}$ parameter. All other parameter errors were also improved, but to a smaller amount. The error for predicting the necessary time to reach a defined moisture level could be reduced significantly by applying the proposed optimization procedure.

In this contribution the benefits of the optimal design of experiment with regard to the mentioned modified Peleg model and a water absorption process of wheat grains is demonstrated. The optimal experimental design offers the opportunity to determine model parameters with a much higher degree of accuracy.

References

Ataíde, F., Hitzmann, B., 2009. When is optimal experimental design advantageous for the analysis of Michaelis-Menten kinetics? Chemometr. Intel. Lab. Syst. 99 (1), 9–18. http://dx.doi.org/10.1016/j.chemolab.2009.07.005.

Bos, A., 2007. Parameter Estimation for Scientists and Engineers. Wiley, Hoboken, NJ, p. 273. http://dx.doi.org/10.1002/9780470173862.

Cunningham, S.E., McMinn, W.A.M., Magee, T.R.A., Richardson, P.S., 2007. Modelling water absorption of pasta during soaking. J. Food Eng. 82 (4), 600–607. http://dx.doi.org/10.1016/j.jfoodeng.2007.03.018.

Delcour, J.A., Hoseney, R.C., 2010. Principles of cereal science and technology, 3rd ed. American Association of Cereal Chemists, St. Paul, Minn. http://dx.doi.org/10.1094/9781891127632.

Dolan, K.D., Mishra, D.K., 2013. Parameter estimation in food science. Ann. Rev. Food Sci. Technol. 4, 401–422. http://dx.doi.org/10.1146/annurev-food-022811-101247.

Erling, P., Botterbrodt, S. (Eds.), 2008. Handbuch Mehl- und Schälmüllerei, 3, überarb., erw. Aufl ed. Agrimedia, Clenze, 522pp. ISBN:978-3860372760.

Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameter precision: state of the art. Model-Based Exp. Anal. 63 (19), 4846–4872. http://dx.doi.org/10.1016/j.ces.2007.11.034.

Maskan, M., 2001. Effect of maturation and processing on water uptake characteristics of wheat. J. Food Eng. 47 (1), 51–57. http://dx.doi.org/10.1016/S0260-8774(00)00099-6.

Munson-McGee, S.H., Mannarswamy, A., Andersen, P.K., 2011a. D-optimal designs for sorption kinetic experiments: cylinders. J. Food Eng. 104 (2), 202–207. http://dx.doi.org/10.1016/j.jfoodeng.2010.12.009.

Munson-McGee, S.H., Mannarswamy, A., Andersen, P.K., 2011b. D-optimal designs for sorption kinetic experiments: slabs. J. Food Eng. 104 (3), 461–466. http://dx.doi.org/10.1016/j.jfoodeng.2011.01.012.

Paquet-Durand, O., Zettel, V., Hitzmann, B., 2014. Optimal experimental design for parameter estimation of the Peleg model. Chem. Intel. Lab. Syst. 140, 36–42. http://dx.doi.org/10.1016/j.chemolab.2014.10.006.

Peleg, M., 1988. An empirical model for the description of moisture sorption curves. J. Food Sci. 53 (4), 1216–1217. http://dx.doi.org/10.1111/j.1365-2621.1988.tb13565.x.

Pukelsheim, F., 1993. Optimal design of experiments. Wiley, New York, p. 454. http://dx.doi.org/10.1137/1.9780898719109.

Sánchez, M.S., Sarabia, L.A., Ortiz, M.C., 2012. On the construction of experimental designs for a given task by jointly optimizing several quality criteria: Pareto-optimal experimental designs. Anal. Chim. Acta 754, 39–46. http://dx.doi.org/10.1016/j.aca.2012.10.014.

Shafaei, S.M., Masoumi, A.A., Roshan, H., 2014. Analysis of water absorption of bean and chickpea during soaking using Peleg model. Journal of the Saudi Society of Agricultural Sciences (in press) (0) http://dx.doi.org/10.1016/j.jssas.2014.08.003.

Sopade, P.A., Ajisegiri, E.S., Badau, M.H., 1992. The use of Peleg's equation to model water absorption in some cereal grains during soaking. J. Food Eng. 15 (4), 269–283. http://dx.doi.org/10.1016/0260-8774(92)90010-4.

Turhan, M., Sayar, S., Gunasekaran, S., 2002. Application of Peleg model to study water absorption in chickpea during soaking. J. Food Eng. 53 (2), 153–159. http://dx.doi.org/10.1016/S0260-8774(01)00152-2.

3.3 A bootstrap based method for optimal design of experiments

By Olivier Paquet-Durand, Viktoria Zettel and Bernd Hitzmann.

Submitted 2015 in **Journal of Chemometrics**.

A bootstrap based method for optimal design of experiments

O. Paquet-Durand*, V. Zettel, B. Hitzmann

*o.paquet-durand@uni-hohenheim.de, Tel: +49 711 459 24630, Fax: +49 711 459 23259

Department of Process Analytics and Cereal Science

Institute of Food Science and Biotechnology

University of Hohenheim

Garbenstr. 23

70599 Stuttgart

Germany

Abstract

Bootstrapping can be used for the estimation of parameter variances and it is straightforward to be implemented but computationally demanding compared to other methods for parameter error estimation. It is not bound to any restrictions such as the distribution of measurement errors. In this work the feasibility of a bootstrap based method for optimal experimental design was evaluated for the Peleg model. First the optimal design was carried out, based on the Cramér-Rao lower bound as a benchmark. Afterwards the optimal design was calculated based on the bootstrap method.

It is demonstrated, that a bootstrap based optimal design of experiments will give comparable results to the Cramér-Rao lower bound optimal designs, however with slightly different measurement points in time and multiplicity. If the parameter errors obtained from both optimal experimental designs are compared, they deviate for the two methods on average by 1.5 %. Due

to the asymmetry of the probability densities of the parameters, the parameter estimation errors acquired by bootstrapping are likely to be more accurate.

Bootstrapping can be used for problems which cannot be solved using Cramér-Rao lower bound due to necessary but invalid assumptions. However, the benefits of the bootstrap method come at the cost of a significant increase in computational effort. The computation time for a bootstrap based optimal design is 25 minutes compared to 5 seconds when using the Cramér-Rao lower bound method. As computers get faster and faster over time, the increase in computational demand will probably become less relevant in future.

Keywords: optimal design of experiments; bootstrapping; Cramér-Rao lower bound; water absorption; Peleg equation

2

1. Introduction

The application of optimal experimental design is increasing almost exponentially during the last years. The basic idea of an optimal experimental design is to optimize the experimental conditions as for example measurement times, locations or number of measurements, to determine parameter values of a mathematical model as accurate as possible as described by Deming et al. [1] and Schlosser [2]. In this context, the parameter estimation errors are usually computed using the Fisher information and Cramér-Rao lower bound (CRlb) [3].

In this work, the feasibility of a bootstrap based method for parameter error estimation in an optimal design procedure was investigated. The mathematical model used here is the Peleg model which is common in food technology for describing water absorption kinetics of many crops [4–6].

Based on the central limit theorem [7], a common assumption made about almost any kind of measurements is that they are normally distributed [8], no matter if they really are or not. This assumption makes calculations a lot easier as mathematical or statistical evaluation methods often require measurement data to be normally distributed. However non-normally distributed measurement data are quite common [9, 10] as many measurement sensor systems will perform an indirect measurement using a mathematical transformation of the raw measurement. Even if the underlying raw physical measurement data might be normally distributed and homoscedastic, the measurement output of the sensor system is not. So this assumption is not conforming to the practical real world and it will probably introduce inaccuracies in subsequent calculations.

Therefore, the need for methods arises, which do not require normality or other restrictive assumptions about the measurement data or errors. For example, Cao et al. [11] used an iterative approach to estimate maximum likelihood for heavy tailed distributions. Another

option can be bootstrapping [12]. The bootstrap method is in fact quite common for parameter error estimation as for example Srivastava and Rawlings [13] and Banks et al. [14] used bootstrapping for estimation of confidence intervals of parameters. Banks even used it with homo- and heteroscedastic data.

Optimal design of experiments based on the Cramér-Rao lower bound is a widespread method. Ataíde and Hitzmann [15] demonstrated, that for a parameter estimation of an enzyme process, a fed-batch design is advantageous compared to a batch design. Bernaerts and Van Impe [16] used optimal experimental design to acquire optimal dynamic model parameters of an *E. coli* cultivation by applying an optimized temperature step during cultivation. Optimal experimental design can also be used to determine detection limits in chemical analysis as shown by Kuselman and Shenhar [17]. Franceschini and Macchietto [18] showed a list of multiple optimal design of experiment applications ranging from chemical kinetics to biological modelling. Martens et al. [19] showed how Monte Carlo simulations can be used to drive an experimental design to determine an ideal size of an experiment.

Arai et al. [20] used a bootstrapping based method for validation of the results of a design of experiments. Gazut et al. [21] used a bootstrapping like resampling method in combination with neurnal networks in an optimal experimental design to reduce prediction variance of the neuronal networks. Therefore, they iteratively added new measurement points in the input space to where the prediction variance (determined by resampling) was maximal.

In this work, the feasibility of using bootstrapping for parameter error estimation to directly drive an optimal experimental design to improve the parameter accuracy of a Peleg model was investigated. The results are compared to the established method based on Cramér-Rao lower bound.

4

2. Theory

The primary goal of optimal experimental design is to determine optimal measurement points and process operation mode (such as batch or fed-batch mode [22]) for an experiment in a way, that the variances of the corresponding model parameters obtained from measurements are as small as possible. For the calculation of the parameter errors or variances of a fitted model, two methods are commonly used: Bootstrapping and the Cramér-Rao lower bound.

The Cramér-Rao lower bound is computationally easy to calculate, but it requires the measurement errors and the parameter errors to be normally distributed. Both requirements are not always fulfilled, therefore in these cases the Cramér-Rao lower bound can be considered to be only a rough estimator for the parameter variance.

The bootstrapping method has no such restrictions on parameter or measurement errors, but it is very demanding from a computation standpoint. It gives no exact solution because the bootstrapping is not deterministic.

2.1 The theoretical model for water absorption

For the water sorption process many different models could be used. A good start would be the Langmuir adsorption isotherm [23]. In this investigation, the empirically found Peleg model [24] (eq. 1), which is basically the same as the Langmuir isotherm but with a slightly more convenient parameter layout, was used.

$$\boldsymbol{f(t) = m_0 + \frac{t}{p_1 + p_2 \cdot t}} \qquad \textbf{1}$$

Here p_1, p_2 and m_0 are the parameters of the model. m_0 is the starting water content of the sample, p_1 is the inverse of the initial water uptake rate and $1/p_2$ determines the final water uptake. If only relative moisturisation is measured as in this work, m_0 can be set to zero and therefore excluded.

5

2.2 Cramér-Rao lower bound and Fisher information

If the measurement errors are normally distributed, the parameter estimation variance can be estimated based on the Cramér-Rao lower bound (eq. 2).

$$\mathbf{var}(\boldsymbol{p_i}) \geq \left[\boldsymbol{I^{-1}}(\vec{\boldsymbol{p}})\right]_{\boldsymbol{ii}} \qquad \mathbf{2}$$

The Cramér-Rao lower bound (CRlb) is a limit of the performance of the best unbiased estimator for the parameter variances, which can be reached under optimal conditions. It is often used to estimate the parameter errors. The lower bound of the estimation error variance of the i^{th} component of parameter vector $\vec{p}$ is determined by using the diagonal elements (marked by $[\dots]_{ii}$) of the inverse (marked by $^{-1}$) of the Fisher information matrix I. The ij-element of the Fisher information matrix is defined in eq. 3:

$$\boldsymbol{I}(\vec{\boldsymbol{p}})_{\boldsymbol{ij}} = -\boldsymbol{E}\left\langle \frac{\boldsymbol{\partial^2 \ln}(\mathbf{pdf}(\boldsymbol{y}, \vec{\boldsymbol{p}}))}{\boldsymbol{\partial p_i \cdot \partial p_j}} \right\rangle \qquad \mathbf{3}$$

Where E<…> is the expectation value of the derivative of the logarithm of the probability density function pdf (…) of the measurement variable y with respect to the model parameters p_i and p_j. The model for the measurement variable is as shown in eq. 4:

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}, \vec{\boldsymbol{p}}) \qquad \mathbf{4}$$

x is the independent variable, like the measurement coordinate or measurement time. The Fisher information matrix is then calculated as follows is eq. 5:

$$\boldsymbol{I}(\vec{\boldsymbol{p}})_{\boldsymbol{ij}} = \sum_{\boldsymbol{k=1}}^{\boldsymbol{N}} \frac{\mathbf{1}}{\boldsymbol{\sigma_k^2}} \frac{\boldsymbol{\partial f}(\boldsymbol{x_k}, \vec{\boldsymbol{p}})}{\boldsymbol{\partial p_i}} \cdot \frac{\boldsymbol{\partial f}(\boldsymbol{x_k}, \vec{\boldsymbol{p}})}{\boldsymbol{\partial p_j}} \qquad \mathbf{5}$$

Here N is the number of measurements and σ_k^2 is the variance of the *k*-th measurement. So the change of the function f with respect to a change of the parameter at the measurement points x_k has an influence on the element of the Fisher information matrix. If at the measurement point x_k the change of $f(x, \vec{p})$ is high with a change of p_i, the information is high, if on the other

hand the function does not depend on the parameter at the measurement point, no additional information is given.

2.3 Bootstrapping method

Another way of estimating parameter variances is the bootstrap method [25]. Bootstrapping is a resampling method, which in principle is easy to implement. It relies on picking random subsamples with replacement of measurement data and fitting the same model to those subsamples over and over again using for example the least squares method. By doing so, from the distribution of the calculated parameter values the second moment can be calculated, whose square root is the estimate of the standard error of the parameter. It does not require any assumption of normality, neither for the measurement data nor for the estimated parameter variation. The probability density functions of both can be asymmetrically and of literally any shape.

Despite its advantages over the Cramér-Rao lower bound estimator until now it has not been used for optimal experimental design and there are also some solid reasons for this:

First, bootstrapping requires a big computational effort for the parameter error estimation. It is usually in the region of some seconds on a todays common desktop computer (2 Ghz Quad Core, 8 GiB RAM) but for an optimization procedure this has to be repeated thousands or even millions of times which quickly requires a huge amount of pure computation time.

Second, bootstrapping is non deterministic. The same parameters and conditions will usually result in slightly different results for the parameter estimation errors. This “statistical noise” is a big problem for optimization and will in fact be problematic for most gradient descent algorithms. To reduce this problem, the resulting parameter variations can be normalized and the shifted Weibull distribution [26] (eq. 6) can be fitted to the results.

7

$$\boldsymbol{pdf_{weibull}(x) = \frac{k}{\lambda} \cdot \left(\frac{x-\mu}{\lambda}\right)^{k-1} \cdot e^{-\left(\frac{x-\mu}{\lambda}\right)^{k}}} \quad \textbf{6}$$

μ is a horizontal shift of the function, λ is a horizontal scale parameter and k is a shape parameter. If k has a value of ~3.6, the Weibull distribution will be symmetric and its shape becomes very close to a normal distribution. The desired percentile values can then be calculated analytically or for an optimal design the value of λ can be used directly as it is usually proportional to the percentiles. This way the computational efforts will be much lower and the statistical noise gets filtered by the fitted Weibull distribution.

2.4 Optimal experimental design

The objective of an optimal experimental design is to optimize a given objective function by variation of the experimental conditions like the measurement points in time $\boldsymbol{x_k}$ [3, 22]. If no a priori information about the experiment is available, an initial experiment is required. For this initial experiment, an equidistant design where measurements are carried out at equidistant measurement points in time (for example every n minutes) is a solid approach. However, the parameter accuracy might be low, because the sensitivity of the function with respect to the parameters and the measurement time points is not considered. Afterwards the optimal experimental design can be computed, using the estimated non accurate parameter values. The corresponding experiments can be carried out as well as the calculation of the parameter values improving their accuracy. If the new estimated parameters deviate very much from the first estimation, the optimal design procedure must be repeated.

In this work, the optimal experimental design was carried out to find the optimal measuring points in time which enable the estimation of the Peleg model parameter as accurate as possible.

8

3. Material and Methods

3.1 Programming and optimization procedure

The programming was carried out on a desktop PC (Intel Core i5-2400, 3100 MHz, 24 GB RAM, MS-Windows 8). The programming and the optimization have been performed using the following software:

- Matlab 2014a (ver. 8.3) as well as the following toolboxes: symbolic math toolbox (ver. 6.0), optimization toolbox (ver. 7.0), parallel computing toolbox (ver. 6.4) and distributed computing toolbox (ver. 6.5).
- C# for the computationally intensive steps (function fitting and bootstrapping)

Because pure gradient descend optimization algorithms will have difficulties solving problems with possibly multiple local minima or statistical noise in the quality function, a genetic algorithm implementation (GA) of Matlab was used to solve the optimization problem. For the GA, the population count was 150. For each optimization up to 150 generations were evaluated. Otherwise the standard settings were used.

3.2 Optimality criterion

For optimal experimental design using the Fisher information and the Cramér-Rao lower bound, a few optimality criteria have been proposed [3]. For a comparison with results of the bootstrap method, an appropriate optimality criterion which can be used with both methods has to be applied. Therefore as criterion the minimum of the sum of relative parameter errors is applied which is called P_r criterion (eq. 7) [27].

$$\boldsymbol{min}\left(\frac{\Delta \boldsymbol{p_1}}{|\boldsymbol{p_1}|}+\frac{\Delta \boldsymbol{p_2}}{|\boldsymbol{p_2}|}\right) \qquad \textbf{7}$$

9

For the Cramér-Rao lower bound method it is the sum of the square root of the diagonal elements of the parameter estimation variance-covariance matrix divided each by the corresponding parameter value. For the bootstrap method it is the sum of the half widths of the 68.2 % confidence intervals divided by the corresponding parameter values.

3.3 Measurement error

For both methods, the estimated measurement errors are required to calculate the parameter errors. In this work, a constant measurement error of 0.1 g_w/g_s (gram water per gram solid) was assumed (eq. 8)

$$\sigma^2 = \left(0.1\frac{g_w}{g_s}\right)^2 = 0.01\frac{g_w^2}{g_s^2} \qquad 8$$

3.4 Experiments

3.4.1 Optimal design of experiments

The goal of this investigation is to find the optimal measurement points in time to get the smallest parameter estimation errors with two different algorithmic approaches and to compare these results.

Some constraints for the experiments and the optimal experimental design have to be defined. In this investigation the maximum measurement time of the experiment is assumed to be 48 hours (2 days). During this time span, 12 measurements are carried out. The parameter values for the Peleg model were p_1=20 $h \cdot g_s/g_w$ and p_2=2 g_s/g_w. Two optimal designs of experiment were calculated:

ODoE-CR: The measurement times were determined by using Cramér-Rao lower bound and minimizing the P_r value.

10

ODoE-BS: The measurement times were determined by using the bootstrap based method to minimize the P_r value.

3.4.2 Parameter distribution analysis

To compare the different design methods (equidistant, based on Cramér-Rao as well as on bootstrap method) the probability density functions of the parameters are calculated. Here the bootstrap method is applied as described in section 3.5.2 using the assumed parameter values as well as the measurement points determined by the design (equidistant, ODoE-CR and ODoE-BS).

3.5 Bootstrap implementation

3.5.1 Bootstrap based method for optimal design of experiments

The bootstrap based optimal design of experiments was implemented as follows: For each individual of the parent genetic algorithm optimization, 50,000 “bootstrap” samples each with 12 measurement points were generated by random sampling with replacement. Then for each sample, the corresponding measurement values were simulated using the Peleg model and the random normally distributed measurement error was added to the simulated value. The Peleg model was fitted to each sample and therefore 50,000 slightly different values for the parameters (p_1 and p_2) were obtained. To the normalized probability density distribution of these parameters a Weibull distribution (eq. 6) was fitted. For the fitting the least squares method combined with a Nelder-Mead simplex algorithm for minimization was used. The parameter errors were then estimated to be the half width of the 0.682 confidence intervals of the fitted Weibull distribution.

3.5.2 Bootstrap based method for the analysis of the parameter distribution

11

For the parameter distribution and parameter error estimation, a similar bootstrap algorithm was used, but the amount of bootstrap samples *m* was increased to 10,000,000. In this case the probability density function was not fitted to the acquired parameter distributions. Instead, the half width of the 0.682 confidence interval was calculated directly by using the bootstrap percentile method.

4. Results & Discussion

4.1 Optimal design of experiments

Figure 1 shows the simulated measurement points without a measurement error of the initial non optimal experiment (equidistant design) of the water uptake and the graph of the corresponding Peleg model. Using the above assumptions, the two optimal experimental designs were carried out. The first one was the classical optimal design of experiment (ODoE-CR) based on the minimization of the Cramér-Rao lower bound estimating the parameter errors. As second approach the bootstrap based method was applied (ODoE-BS). The computation time for ODoE-CR was about 5 seconds, for ODoE-BS it was about 25 minutes on the aforementioned desktop computer. The resulting measurement points of both optimal design procedures are shown in Figure 1 as well.

As the Peleg model has 2 parameters, there must be 2 optimal measurement time points. Both ODoE runs resulted in 2 optimal points in slightly different repetitions, if the average of the two blocks from ODoE-BS are used. As described by Paquet-Durand et al. [27], the first optimal measurement point can be estimated to be approximately 5.9 h. The exact values of the two sets of computed optimal measurement points as well as the averages of the two expected optimal measurement points are presented in Table 1.

For the ODoE-CR experiment, the values of the calculated measurement time points are $t_{1,CR}$=5.98 h as well as $t_{2,CR}$=48 h with the multiplicity of 7 and 5 respectively. The ODoE-BS

measurement points are scattered around the two ideal measurement points, the corresponding average values are t_{1BS}=5.13 h and $t_{2,BS}$=46.6 h. However, as one can see, both points in time are smaller for ODoE-BS experiment. The multiplicities of ODoE-BS show the same trend as the one of ODoE-CR but here 8 and 4 are determined. The likely reason for this deviation is assumed to be the aforementioned (section 2.3) statistical noise problem within the bootstrap based ODoE-BS results. As the optimization algorithm converges towards the optimal measurement points, this noise will limit further convergence. So the "optimal" measurement points achieved by ODoE-BS are just an approximation. The quality of this approximation depends on the amount of statistical noise and therefore on the amount of bootstrap samples or the assumed measurement error.

In case of the ODoE-CR method, the measurement error (when constant) is in theory irrelevant for the values of the optimal measurement points. A different measurement error would change the magnitude of the corresponding sensitivities and therefore the parameter estimation errors. But the qualitative shape of the sensitivities will still be the same. The position of the optimal measurement point would not change. For the ODoE-BS method on the other hand higher assumed measurement errors will result in a broader parameter distribution and therefore a higher statistical noise. This will then reduce the convergence speed of the overlying optimization algorithm and might as well change the outcome of the optimization.

4.2 Parameter estimation errors and parameter distribution analysis

In Figure 2 the probability density functions of the parameters are presented as well as the assumed parameter values. As one can see, for all distributions the assumed parameter values are near the maximum values and the probability densities of the parameters are neither normally distributed nor even symmetrical. The expected values of all distributions are the corresponding assumed parameter values. The estimated probability density functions of the

13

parameters p_1 and p_2 were calculated by bootstrapping as described above. It is obvious, that the equidistant design gave the broadest distribution. There are just small deviations comparing the distributions obtained by ODoE-CR and ODoE-BS.

In Figure 2 it can be seen, that all three experiments resulted in an asymmetric probability density for the parameter p_1. For p_2 the probability densities of the parameter for all experiments are almost symmetrical. To confirm the asymmetry, the skewness values for these probability densities were calculated and are presented in Table 2.

Skewness values bigger than 1 or smaller than -1 are usually considered as an indication for a significant asymmetry. As stated above, the probability densities for p_1 are all in fact significantly asymmetric and for p_2 the asymmetries are not significant. A Jarque-Bera test [28] for normality was also applied to the acquired distributions. The resulting values for the test ranged from 8.4 * 10^5 to 5.0 * 10^6. Assuming a significance level of 0.01 the critical value would be 9.2. Therefore, it is safe to assume that the distributions are not normally distributed. The extreme high values are caused by the very high number of bootstrap samples which is a linear factor in the calculation of the Jarque-Bera test.

For all experiments, the parameter errors were estimated using both methods, the Cramér-Rao lower bound and bootstrapping method as described above. These results can be seen in Table 3.

The parameter estimation error for the first non-optimal experiment calculated based on CRlb is 57.3 % for p_1. For p_2 the estimation error is 22.5 %. The reason for the error of p_1 being so big is as follows: The initial water uptake rate represented by p_1, is of course determined early in the process. When regarding the optimal design results, the ideal time for the determination is at about 5 – 6 h. The closest measurement in the non-optimal experiment was 4 h which is quite

far away. Using the bootstrap method to determine the errors, the trend is the same ($\Delta p_1 > \Delta p_2$), however the values are higher.

Both optimal design methods on the other hand resulted in much smaller parameter estimation errors. Using ODoE-CR and CRlb the error for p_1 is 38 % and p_2 is 17 %. This demonstrates the ability of optimal experimental design. If the ODoE-CR experiment is analysed using the bootstrap method, the errors are increasing slightly. The error of p_1 calculated with the measurement points of ODoE-CR and CRlb is slightly higher, than the corresponding value calculated by ODoE-BS and CRlb. This is a result of $t_{1,BS}$ is smaller than $t_{1,CR}$ as well as the higher multiplicity of $t_{1,BS}$. For the same reason the corresponding errors of p_2 are higher.

It is obvious that the differences in the parameter estimation errors for the ODoE-CR and ODoE-BS experiment are very small in comparison to the value of the estimated parameter error. The average parameter error is 27.5 % for ODoE-CR and 27.7 % for ODoE-BS calculated by CRlb method and 32.9 % and 33.2% respectively calculated by bootstrap method. The parameter errors for ODoE-CR and ODoE-BS deviate by 1.5 % on average. The reduction of the parameter estimation errors with respect to the equidistant design is for both optimal design methods comparable.

Due to the asymmetry of the probability densities of the parameter errors, the calculated parameter errors using the bootstrap method are bigger than the errors estimated by the Cramér-Rao lower bound. This is to be expected, as the Cramér-Rao lower bound is, as the name supposes, the lowest value of the parameter estimation error, which is only valid under optimal conditions and if the probability density is symmetric and normally distributed. Both is not the case, so the actual parameter estimation errors are expected to be bigger than the Cramér-Rao lower bound.

15

5. Conclusion

In this contribution the utilization of a bootstrap based method for the optimization part of optimal design of experiments is presented and compared to the classical method based on the Cramér-Rao lower bound. Both methods provide comparable results for the optimal designed experiment.

Using the bootstrap based method required significantly more computation time of about 25 minutes, compared to 5 seconds for the CRlb method. But the computational demand for the bootstrap method could be reduced significantly from days or weeks to minutes by fitting an appropriate probability density function to the distribution of the parameter values which were obtained from the bootstrap method. In this case the Weibull distribution was used. The reduction of computational time was significant and enables the bootstrap method to be used for the optimization part of an optimal design of experiments in an acceptable period of time. As the available computing power increases further over time, the necessary computational effort will probably become less significant in the foreseeable future.

By using the bootstrapping like resampling method for parameter variation estimation, not only the parameter estimation errors can be calculated more accurately but additionally, information about the real probability density of the parameters distribution can be gathered. And in contrast to the CRlb method, the bootstrap like resampling method can be applied to scenarios with any measurement distribution. For CRlb method normally distributed measurement and parameter variances have to be assumed although this rarely matches the real world. Therefore, large errors can be introduced to the parameter error estimation and to the optimal design of experiments. In this cases, the CRlb method should not be applied for the determination of the parameter estimation errors as well as for optimal design of experiments. The bootstrap method however can be applied without any restriction.

The main problem applying bootstrapping for optimal design of experiments is, that in contrast to the CRlb method, the bootstrapping results are non-deterministic and will contain statistical noise. This drastically reduces the convergence speed of most optimization algorithms. By using a more robust optimization algorithm one might improve the bootstrap based optimal design of experiments in the future.

References

1 Deming, S.N., Palasota, J.A., Palasota, J.M. (1991) Experimental design in chemometrics. *Journal of Chemometrics*, **5** (3), 181–192.

2 Schlosser, P.M. (1994) Experimental design for parameter estimation through sensitivity analysis. *Journal of toxicology and environmental health*, **43** (4), 495–530.

3 Pukelsheim, F. (1993) *Optimal design of experiments*, Wiley, New York.

4 Cunningham, S.E., McMinn, W. A. M., Magee, T. R. A., Richardson, P.S. (2007) Modelling water absorption of pasta during soaking. *Journal of Food Engineering*, **82** (4), 600–607.

5 Maskan, M. (2001) Effect of maturation and processing on water uptake characteristics of wheat. *Journal of Food Engineering*, **47** (1), 51–57.

6 Sopade, P.A., Ajisegiri, E.S., Badau, M.H. (1992) The use of Peleg's equation to model water absorption in some cereal grains during soaking. *Journal of Food Engineering*, **15** (4), 269–283.

7 Fischer, H. (2011) *A history of the central limit theorem: From classical to modern probability theory*, Springer, New York, NY.

8 Jaynes, E.T. (2003) *Probability theory: The logic of science*, Cambridge University Press, Cambridge, UK, New York, NY.

9 Pyzde, T. (1995) Why normal distributions aren't [all that normal]. *Quality Engineering*, **7** (4), 769–777.

10 CLANCEY, V.J. (1947) Statistical Methods in Chemical Analyses. *Nature*, **159** (4036), 339–340.

11 Cao, C., Lin, J., Shi, J.Q., Wang, W., Zhang, X. (2015) Multivariate measurement error models for replicated data under heavy-tailed distributions. *Journal of Chemometrics*, **29** (8), 457–466.

12 Ortiz, M.C., Sarabia, L.A., Sánchez, M.S. (2010) Tutorial on evaluation of type I and type II errors in chemical analyses: From the analytical detection to authentication of products and process control. *Analytica Chimica Acta*, **674** (2), 123–142.

17

13 Srivastava, R. and Rawlings, J.B. (2014) Parameter estimation in stochastic chemical kinetic models using derivative free optimization and bootstrapping. *Computers & Chemical Engineering*, **63** (0), 152–158.

14 Banks, H.T., Holm, K., Robbins, D. (2010) Standard error computations for uncertainty quantification in inverse problems: Asymptotic theory vs. bootstrapping. *Mathematical and Computer Modelling*, **52** (9–10), 1610–1625.

15 Ataíde, F. and Hitzmann, B. (2009) When is optimal experimental design advantageous for the analysis of Michaelis–Menten kinetics? *Chemometrics and Intelligent Laboratory Systems*, **99** (1), 9–18.

16 Bernaerts, K. and Van Impe, J. (2005) Optimal dynamic experiment design for estimation of microbial growth kinetics at sub-optimal temperatures: Modes of implementation. *Simulation Modelling Practice and Theory*, **13** (2), 129–138.

17 Kuselman, I. and Shenhar, A. (1995) Design of experiments for the determination of the detection limit in chemical analysis. *Analytica Chimica Acta*, **306** (2–3), 301–305.

18 Franceschini, G. and Macchietto, S. (2008) Model-based design of experiments for parameter precision: State of the art. *Model-Based Experimental Analysis*, **63** (19), 4846–4872.

19 Martens, H., Dijksterhuis, G.B., Byrne, D.V. (2000) Power of experimental designs, estimated by Monte Carlo simulation. *Journal of Chemometrics*, **14** (5-6), 441–462.

20 Arai, H., Suzuki, T., Kaseda, C., Takayama, K. (2009) Effect of an experimental design for evaluating the nonlinear optimal formulation of theophylline tablets using a bootstrap resampling technique. *Chemical & pharmaceutical bulletin*, **57** (6), 572–579.

21 Gazut, S., Martinez, J.-M., Dreyfus, G., Oussar, Y. (2008) Towards the optimal design of numerical experiments. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, **19** (5), 874–882.

22 Lindner, P.F.O. and Hitzmann, B. (2006) Experimental design for optimal parameter estimation of an enzyme kinetic process based on the analysis of the Fisher information matrix. *Journal of Theoretical Biology*, **238** (1), 111–123.

23 Langmuir, I. (1933) *Surface chemistry: Nobel lecture, presented in Stockholm on December 14, 1932*, Stockholm.

24 Peleg, M. (1988) An empirical model for the description of moisture sorption curves. *Journal of Food Science*, **53** (4), 1216–1217.

25 Efron, B. and Tibshirani, R.J. (1998) *An introduction to the bootstrap*, 1st edn, Chapman & Hall/CRC, Boca Raton Fla. u.a.

26 Johnson, N.L., Kotz, S., Balakrishnan, N. (1994) *Continuous univariate distributions Vol. 1*, 2nd edn, Wiley, New York.

18

27 Paquet-Durand, O., Zettel, V., Hitzmann, B. (2014) Optimal experimental design for parameter estimation of the Peleg model. *Chemometrics and Intelligent Laboratory Systems*, 36–42.

28 Jarque, C.M. and Bera, A.K. (1980) Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, **6** (3), 255–259.

Appendix B: Figures

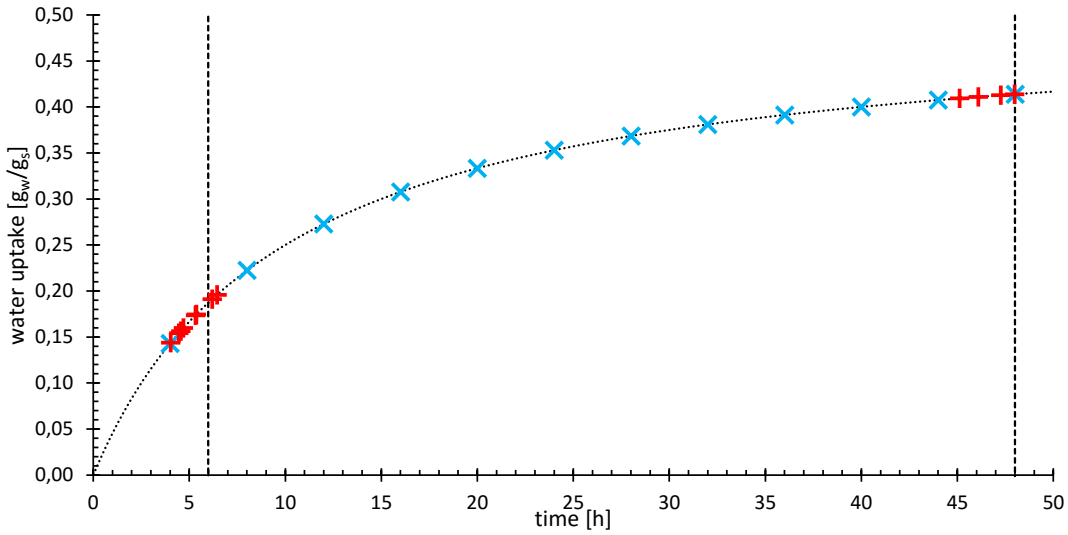

Figure 1: Simulated Measurement points without a measurement error; the corresponding measurement time points are obtained from the three designs; blue symbols: equidistant design, vertical dotted lines: optimal design using CRlb method (ODoE-CR), red symbols: optimal design result using bootstrap based method (ODoE-BS).

1

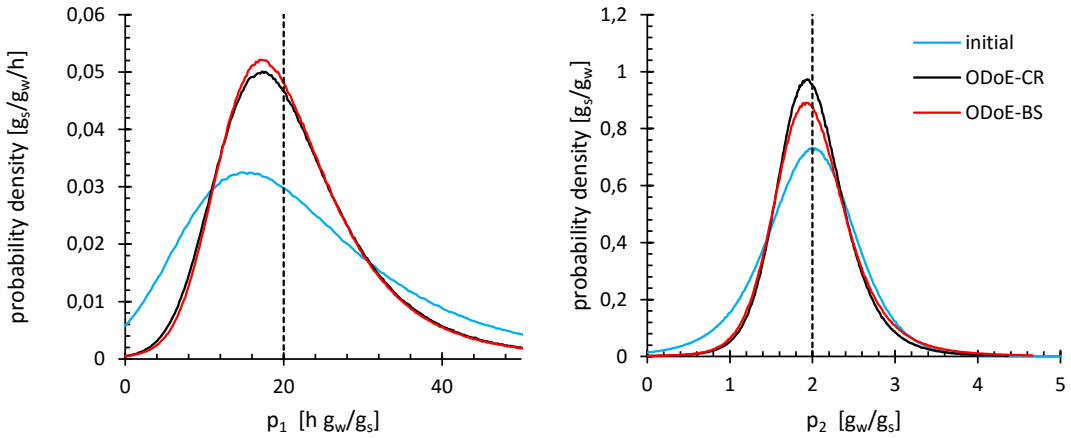

Figure 2: Estimated parameter probability densities of the three experiments, acquired by normalization of the parameter values histograms obtained by the bootstrapping method. The vertical dotted line shows the assumed parameter values, which are also the average values of the corresponding distributions.

2

Appendix A: Tables

Table 1: Resulting measurement time points of the two optimal design approaches as well as the average values for the two expected optimal points.

ODoE-CR [h]	5.98	5.98	5.98	5.98	5.98	5.98	5.98	48.00	48.00	48.00	48.00	48.00
	7x 5.98 ± 0.00							5x 48 ± 0.00				
ODoE-BS [h]	4.03	4.45	4.55	4.69	5.32	5.36	6.19	6.44	45.11	46.08	47.25	47.97
	8x 5.13 ± 0.96							4x 46.60 ± 1.26				

Table 2: Skewness values for the probability densities of the three experiments.

	equidistant	ODoE-CR	ODoE-BS
p_1	1.11	1.191	1.116
p_2	-0.312	0.465	0.474

Table 3: Parameter estimation errors calculated by Cramer Rao lower bond method and by bootstrapping of simulated values for the three experiments.

	equidistant		ODoE-CR		ODoE-BS	
	CRlb	Bstrp.	CRlb	Bstrp.	CRlb	Bstrp.
Δp_1 [h·g_w/g_s]	11.45 (57.3%)	14.00 (70.0 %)	7.59 (38.0 %)	8.85 (44.3 %)	7.36 (36.8 %)	8.57 (42.9 %)
Δp_2 [g_w/g_s]	0.45 (22.5 %)	0.57 (28.5 %)	0.34 (17.0 %)	0.43 (21.5 %)	0.37 (18.5%)	0.47 (23.5 %)

1

Chapter 4

Conclusion and final remarks

4.1 Conclusion

Experimental design is a very important field in science. The mathematical foundation was laid by Ronald Fisher in the 1920s and 1930s (Fisher, 1935). Optimal experimental design is a special case of optimal design of experiments, which was developed in the late 1960s (Lyman Ott and Myers, 1968; Stigler, 1971).

Before the milling of cereal grains, there is usually a tempering step. The tempering is very important, because it simplifies the sifting of mill stocks and flour and can reduce power consumption of roller mills as well (Fang and Campbell, 2001; Klingler, 2010). It consists of two steps, the hydration and a resting phase (Hsu, 1984).

To be able to optimize the hydration step, an appropriate theoretical model with accurate parameter values must be known. The work in this contribution is based on the Peleg model. For the determination of as accurate as possible parameter values, an optimal design of experiments has been evaluated. Multiple optimality criteria were investigated for their suitability and the resulting parameter estimation errors were compared to each other and to the resulting parameter errors of non-optimal measurements. These results are summarized in Table 4.1.

Table 4.1: Effects of different optimality criteria in optimal design of experiments on the two parameter errors of the Peleg model.

Assumed parameter value		Optimality criterion / Parameter estimation error						
			equidst	log equidst.	A	D	E	P_r
p_1	$0.5\ min \cdot g_s/g_w$	Δp_1	21.6 %	13.0 %	8.2 %	9.8 %	8.2 %	8.4 %
p_2	$0.05\ g_s/g_w$	Δp_2	3.2 %	4.2 %	4.2 %	2.6 %	4.2 %	3.4 %
		$\sum \Delta p$	24.8 %	17.2 %	12.4 %	12.4 %	12.4 %	11.8 %

It is quite clear that the reduction of the parameter estimation error Δp_1 leaves much more room for improvement than Δp_2. Therefore, it is not surprising that the estimation error Δp_1 of the Peleg model could be improved significantly by up to 62 % (A- and E-optimality) whereas for the second parameter, the estimation error could only be reduced by 19 % (D-optimality). In all but one cases the Δp_2 was slightly bigger when applying an optimal design. This behaviour was to be expected as the optimization algorithm trades a little increase in Δp_2 for a significant decrease in Δp_1. This is a basic problem when trying to solve a multi-objective optimization where multiple parameter estimation errors should be minimized at the same time (Zeleny, 1974). If an unbiased and unweighted optimality should be achieved, the D-optimality can be

used (Pukelsheim, 1993). The lowest sum of all errors is achieved by the P_r criterion as the sum of relative errors is actually the criterion (Paquet-Durand *et al.*, 2015).

In case of the ordinary Peleg equation and the D-optimality, a solution for the optimal experimental design was analytically calculated.

$$t_1 = f(t_2, \vec{p}) = \frac{t_2}{2+\frac{p_2}{p_1} \cdot t_2}$$

With this function the optimal first measurement point in time t_1 can be calculated directly as a function of the final measurement time t_2 and the initial parameter guesses $\vec{p}$. The relatively tedious optimization, which is usually time consuming is not necessary. With the knowledge of this exact solution for the D-optimality, a more general approximation function for the A-, E- and P_r criterion was developed by repeatedly performing the normal optimal design procedure while varying t_2. The following function was found to describe the development of t_1 well:

$$t_1 = f(t_2, \vec{p}, a) \approx \frac{t_2}{a \cdot (1+\frac{1}{a}+\frac{p_2}{p_1} \cdot t_2)}$$

This approximation incorporates an additional parameter a which depends on the optimality criterion. For the D-Criterion, a equals 1, for the A-, E- and P_r-Criterion a is ≥ 1.

The differences in the resulting parameter estimation errors between a normal optimal experimental design using a genetic algorithm and the presented approximation function are negligible and zero for D-optimality.

For a more accurate description of the hydration kinetics of wheat grains, a more appropriate model was derived from the Peleg model by adding another Peleg like term:

$$f(t, \vec{p}) = m_0 + \frac{t}{p_1+p_2 \cdot t} + \frac{t}{p_3+p_4 \cdot t}$$

Subsequently an optimal design of experiments has been evaluated, based on this modified Peleg model. Two experiments had to be carried out to determine the water absorption kinetics of wheat grains. In the first initial experiment 12 measurements were performed and the measurement times were chosen based on experience to acquire rough parameter values. In the second experiment again 12 measurements were carried out but this time the measurement points were calculated by optimal design based on the Cramér-Rao lower bond. The parameter estimation errors for the modified Peleg equation could be significantly reduced by a factor of up to 17.4.

It has been shown that the modified Peleg model is more suitable to describe the hydration

process of wheat grains, whereas the ordinary Peleg model was comparatively inaccurate. The prediction error was reduced from 2.51 % for Peleg model to 0.29 % for the modified Peleg model. However, the number of parameters has been doubled. The error for the prediction of the time required to reach a certain moisture level could also be reduced significantly. The time needed to reach for example a water uptake of 10 % is about 12 minutes. The error for this time estimation is about ± 30 minutes for the parameters acquired with the non-optimal experiment and about ± 6 minutes for the parameters acquired with the optimally designed experiment.

Another possible improvement to the optimal design of experiments is presented in chapter 3.3 where a bootstrap based method was used for parameter error estimation instead of the Cramér-Rao lower bound. The feasibility of using this bootstrap method in optimal design of experiments was evaluated and the results were compared to the classical method. By using the bootstrap based method the computational requirement is increasing by a huge amount from a few seconds to several weeks or months. The main problem was that bootstrapping is non-deterministic and the results will contain statistical noise. An approach to address this issue as well as reduce the amount of required bootstrap samples was the fitting of an appropriate probability density function. In this case a Weibull distribution was used, to approximate the distribution of the parameter values. By doing so, the required bootstrap sample size as well as amount of statistical noise could be reduced significantly. The lesser noise also increased the efficiency or convergence speed of the optimization algorithm. Therefore the computation time could be reduced to about 25 minutes for a complete optimal design of experiments computation. 25 minutes is a reasonable computation time. Especially when considering, that the water absorption experiment itself took 24 to 48 hours. So the bootstrap method can be used for optimal design of experiments in an acceptable period of time.

As it can be expected, that the available computing power will increases further in the near future, the additional computational effort required will probably become negligible. The major benefit of a bootstrap based method for parameter error estimation is, that it can be applied to scenarios with any measurement or parameter distribution without any restriction. The Cramér-Rao lower bound requires the parameter distributions to be normally distributed. However, this rarely matches the real world. So, quite significant errors can be introduced to the parameter error estimation and therefore to the optimal design of experiments as well.

4.2 Final remarks

There is usually a trade-off to make between accuracy or quality of the results on one side and complexity or computation time on the other side. The same is obviously true for optimal design of experiments as the optimization in the optimal design process is usually time consuming and computationally demanding. So cost-effectiveness considerations should be made, not only after each iteration of an optimal design as the cost or effort increases but also when considering which methods and algorithms should be used.

Bootstrapping can enable one to apply optimal experimental design in cases where this was not accurately possible before by eliminating systematic errors through avoiding false assumptions. But the computational cost for bootstrapping is very high compared to Cramér-Rao lower bound based optimal design.

On the other hand, an approximation function for optimal design has also been shown to work quite well. With this approximation accuracy could be sacrificed for a massive reduction in required computation time from a few seconds to a few clock cycles of a CPU translating to a couple of nanoseconds.

It has to be decided whether accuracy or fast computation time should be prioritized. In this work, the optimal design procedure has been applied and two potential improvements for either desire were suggested. These improvements should not be seen as all the better ways to go, but as additional options to strike the desired balance between accuracy and complexity.

Bibliography

Arai, H., Suzuki, T., Kaseda, C., Takayama, K., 2009. Effect of an experimental design for evaluating the nonlinear optimal formulation of theophylline tablets using a bootstrap resampling technique. Chemical & pharmaceutical bulletin 57 (6), 572–579.

Ataíde, F., Hitzmann, B., 2009. When is optimal experimental design advantageous for the analysis of michaelis–menten kinetics? Chemometrics and Intelligent Laboratory Systems 99 (1), 9–18.

Banks, H. T., Holm, K., Robbins, D., 2010. Standard error computations for uncertainty quantification in inverse problems: Asymptotic theory vs. bootstrapping. Mathematical and Computer Modelling 52 (9–10), 1610–1625.

Butcher, J., Stenvert, N. L., 1973. Conditioning studies on australian wheat. i. the effect of conditioning on milling behaviour. Journal of the Science of Food and Agriculture 24 (9), 1055–1066.

Cramér, H., 1946. Mathematical Methods of Statistics. Princeton Univ. Press, Princeton, NJ.

Efron, B., Tibshirani, R. J., 1998. An introduction to the bootstrap, 1st Edition. Vol. 57 of Monographs on statistics and applied probability. Chapman & Hall/CRC, Boca Raton Fla. u.a.

Fang, C., Campbell, G. M., 2001. On predicting roller milling performance v: Effect of moisture content on the particle size distribution from first break milling of wheat. Journal of Cereal Science 37 (1), 31–41.

Fisher, R. A., 1935. The Design of Experiments. Oliver and Boyd, Edinburgh, UK.

Fraser, A. S., 1957. Simulation of genetic systems by automatic digital computers. 1. introduction. Australian Journal of Biological Sciences 10 (1), 484–491.

Hsu, K. H., 1984. A theoretical approach to the tempering of grains. Cereal Chemistry 61 (5), 466–470.

Kay, M., 2013. Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory. Prentice Hall PTR, NJ.

Klingler, R. W., 2010. Grundlagen der Getreidetechnologie. B. Behr's Verlag GmbH, Hamburg.

Lindner, P. F. O., Hitzmann, B., 2006. Experimental design for optimal parameter estimation of an enzyme kinetic process based on the analysis of the fisher information matrix. Journal of Theoretical Biology 238 (1), 111–123.

Lyman Ott, R., Myers, R. H., 1968. Optimal experimental designs for estimating the independent variable in regression. Technometrics 10 (4), 811–823.

Montanuci, F. D., Perussello, C. A., de Matos Jorge, Luiz Mario, Jorge, Regina Maria Matos, 2014. Experimental analysis and finite element simulation of the hydration process of barley grains. Journal of Food Engineering 131 (0), 44–49.

Nelder, J. a., Mead, R., 1965. A simplex method for function minimization. Computer Journal 7 (1), 308–313.

Paquet-Durand, O., Zettel, V., Hitzmann, B., 2015. Optimal experimental design for parameter estimation of the peleg model. Chemometrics and Intelligent Laboratory Systems 140, 36–42.

Peleg, M., 1988. An empirical model for the description of moisture sorption curves. Journal of Food Science 53 (4), 1216–1217.

Piergiovanni, A. R., 2011. Kinetic of water adsorption in common bean: Considerations on the suitability of peleg's model for describing bean hydration. Journal of Food Processing and Preservation 35 (4), 447–452.

Pukelsheim, F., 1993. Optimal design of experiments. Wiley series in probability and mathematical statistics Probability and mathematical statistics. Wiley, NY.

Pyzde, T., 1995. Why normal distributions aren't [all that normal]. Quality Engineering 7 (4), 769–777.

Rao, C. R., 1945. Information and the accuracy attainable in the estimation of statistical parameters. Bulletin of the Calcutta Mathematical Society 37 (1), 81–89.

Shewry, P. R., 2009. Wheat: Chemistry and Technology, Fourth Edition. Amer Assn of Cereal Chemists.

Sopade, P. A., Ajisegiri, E. S., Badau, M. H., 1992. The use of peleg's equation to model water absorption in some cereal grains during soaking. Journal of Food Engineering 15 (4), 269–283.

Srivastava, R., Rawlings, J. B., 2014. Parameter estimation in stochastic chemical kinetic models using derivative free optimization and bootstrapping. Computers & Chemical Engineering 63 (0), 152–158.

Stigler, S. M., 1971. Optimal experimental designs for polynomial regression. Journal of the American Statistical Association 66 (334), 311–318.

Turhan, M., Sayar, S., Gunasekaran, S., 2002. Application of peleg model to study water absorption in chickpea during soaking. Journal of Food Engineering 53 (2), 153–159.

Van den Bos, A., 2007. Parameter estimation for scientists and engineers. Wiley, Hoboken, NJ.

Zeleny, M., 1974. A concept of compromise solutions and the method of the displaced ideal. Computers and Operations Research 1, 479–496.

Appendices

A Eidesstattliche Versicherung

UNIVERSITÄT HOHENHEIM
FAKULTÄT NATURWISSENSCHAFTEN

Anlage 2 zur Promotionsordnung der Universität Hohenheim zum Dr. rer. nat.

Eidesstattliche Versicherung gemäß § 7 Absatz 7 der Promotionsordnung der Universität Hohenheim zum Dr. rer. nat.

1. Bei der eingereichten Dissertation zum Thema

   ……..

   ……..

   handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Ich habe nicht die Hilfe einer kommerziellen Promotionsvermittlung oder -beratung in Anspruch genommen.

4. Die Bedeutung der eidesstattlichen Versicherung und der strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Die Richtigkeit der vorstehenden Erklärung bestätige ich: Ich versichere an Eides Statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

_____ _____

Ort und Datum Unterschrift