Wheat baking quality in a process analytical perspective

Sampling, diversification, prediction and chemometric method development

PhD thesis by
Erik Tønning

FACULTY OF LIFE SCIENCES
UNIVERSITY OF COPENHAGEN

Faculty of Agricultural Sciences
UNIVERSITY OF AARHUS

Wheat baking quality in a process analytical perspective

Sampling, diversification, prediction and chemometric method development

PhD thesis by
Erik Tønning, MSc Environmental Chemistry

Frederiksberg, 2007

Supervisors:
Associate Professor Lars Nørgaard
Professor Søren Balling Engelsen

Head of research unit
Anette Kistrup Thybo

Quality & Technology Research Group
Department of Food Science
Faculty of Life Sciences
University of Copenhagen
Rolighedsvej 30
DK-1958 Frederiksberg C
Denmark
www.models.life.ku.dk

Plant Food Science Group
Department of Food Science
Faculty of Agricultural Sciences
University of Aarhus
Kirstinebjergvej 10
DK-5792 Aarslev
Denmark
www.agrsci.org

LMC – IMPROVING FOOD

FOOD GRADUATE SCHOOL

Cover illustration by Michelle Claire Ebbesen
Title: Wheat baking quality in a process analytical perspective

ISBN: 87-91949-23-8
Printed by DigiSource Danmark A/S, Viborg, Denmark

Preface

This PhD thesis finalises three years exploration of wheat quality, spectroscopy and chemometrics at Faculty of Life Sciences (LIFE), University of Copenhagen (KU) and Faculty of Agricultural Sciences (DJF), University of Aarhus (AU). The work was primarily financed through a PhD scholarship from The Royal Veterinary and Agricultural University, now known as LIFE, KU. The project was also on the budget and hosted by The Danish Institute of Agricultural Sciences, now known as DJF, AU. A number of other collaborators also took upon them a significant burden by generously offering material, time, space and equipment for which they are greatly acknowledged: ACABS, Aalborg University Esbjerg, Sejet Planteforædling, Vejle Mølle, BoMill AB, Crop & Food Research, and Weston Milling.

Special gratitude is shown to Paul T. Callaghan and all his staff at The MacDiarmid Institute, Victoria University of Wellington, New Zealand for hosting my crash education in applied nuclear magnetic resonance and for a memorable time with the NMR-gang. Travel grants from Q-interline A/S, Knud Højgaards Fond and Danmarks Jordbrugsvidenskabelige Ph.D.-forening facilitated six extraordinary months in Wellington.

I am deeply thankful to *all* my colleagues (former, present and new) at Quality and Technology, LIFE and Plant Food Science, DJF with whom the time is always well spent both professionally and socially, whatever the difference. I am indebted to my supervisors Lars Nørgaard, Søren B. Engelsen and Anette K. Thybo. Without their confidence, encouragement and strong support this project would not have come this far. I am thankful for the inspiration and supportive engagement by Kim H. Esbensen, Lene Pedersen, Bernd Wollenweber, Johannes R. Jørgensen and Lars Munck. Bo Löfqvist and Daniel Polders, how can I ever thank you for your temporal and spatial timely presence and engagement. Special thanks to Mogens Pedersen, Helle I. Høgild, Jette R. Nielsen and Lisbeth T. Hansen. I am especially grateful for the patience, sympathy and support from Signe, Jakob and Michelle as well as from my friends and family during this quest.

Erik Tønning, Frederiksberg, August 2007

Abstract

The complex and mysterious conversion of grain into flour, visco-elastic dough and soft spongy bread crumb has been studied from three different process analytical technology (PAT) perspectives.

I. The heterogeneity of bulk wheat with respect to protein content was studied in order to quantify the variances and biases of the sampling process errors generated according to the theory of sampling (TOS). The analysis of individual grains in various lots of wheat showed a great variation in mean protein contents and heterogeneity. This inherent heterogeneity was shown to control the variances of the fundamental sampling errors (FSE) and the grouping and segregation errors (GSE) obtained in composite samples. The variances of the sampling errors were an order of magnitude higher than the variances of the total analytical errors (TAE). The sampling process was shown to influence the variances and biases of the results; hence common grab sampling was shown to generate unrealistic variance estimates and biased results as opposed to representative sampling by riffle splitting (Paper I).

II. A holistic view on the entire process from grain to bread was applied in wheat material diversified by agronomical measures as well as novel post harvest single-kernel near-infrared (SKNIR) sorting utilising the inherent constitutional heterogeneity of the internal complex quality traits. The effects of SKNIR sorting and agronomic treatments were quantified and compared and the functionality of the flour as well as the end product quality was predicted from multivariate spectroscopic analysis of grain, flour and dough. The SKNIR fractionation had significant effect on several protein and α-amylase activity related parameters measured. Hence, the flour protein content was increased by 0.4 to 1.7%-points, wet gluten content was increased by 1.8 to 5.5%-points, Zeleny sedimentation volume was increased by 1.4 to 3.5 mL, Farinograph water absorption was increased by 0.5 to 1.4%-points and falling number was increased by 10 to 48 s in the best of three equally sized fractions as compared to the starting materials (Paper II). The prediction of flour functionality with partial least squares projections to latent

structures (PLS) and multi-block PLS (MBPLS) modelling was best using near-infrared reflection (NIR) spectroscopy of flour followed by near-infrared transmission (NIT) spectroscopy of grain and flour, infrared (IR) spectroscopy and nuclear magnetic resonance (NMR) baking relaxometry and combinations thereof. The flour functionality parameters themselves were superior for prediction of corresponding bread quality as compared to the above mentioned spectroscopic methods due to unique information in the functionality parameters not well modelled by spectroscopy in the first place (Paper III).

III. A novel chemometric method, 2D PARAFAC-Laplace decomposition, was developed for unique resolvation, quantification and interpretation of 2-dimensional diffusion-relaxation NMR data obtained in bread dough of fat and water compartmentalisation and dynamics (Paper IV).

Resumé

Den komplekse og forunderlige omdannelse af korn til mel, viskoelastisk dej og blød porøs brødkrumme er blevet undersøgt fra tre forskellige indfaldsvinkler baseret på proces analytisk teknologi (PAT).

I. Proteinheterogeniteten i hvedepartier blev undersøgt for at kvantificere varianser og systematiske fejl på afvigelser genereret i prøvetagningsprocessen jævnfør teorien om sampling (TOS). Proteinanalyser af enkeltkerner i forskellige partier viste en stor variation i gennemsnitsindhold og heterogenitet. Det blev vist, at den iboende heterogenitet styrer de opnåede varianser af fundamentale prøvetagningsfejl (FSE) og grupperings- og segregeringsfejl (GSE) ved målinger af sammensatte prøver. Varianserne på prøvetagningsfejlene var en størrelsesorden større end varianserne på de totale analysefejl (TAE). Desuden blev det vist, at prøvetagningsprocessen havde betydning for varianser og systematiske fejl i resultaterne, idet almindelig prøvetagning med ske gav urealistiske variansestimater og resultater med systematiske fejl i modsætning til repræsentativ prøvetagning med spaltedeler (Artikel I).

II. Der blev anlagt et holistisk perspektiv på hele processen fra korn til brød. Hvedematerialet blev differentieret agronomisk samt ved fraktionering efter høst ved hjælp af en ny nærinfrarød enkeltkernesortering (SKNIR), som udnytter materialets iboende konstitutionelle heterogenitet med hensyn til kernernes indre kvalitet. Effekten af SKNIR-sorteringen og agronomiske behandlinger blev kvantificeret og sammenlignet. Funktionaliteten af melet og kvaliteten af det færdige produkt blev prædikteret med multivariat spektroskopisk analyse af korn, mel, dej og brød. SKNIR-fraktioneringen havde signifikant effekt på flere målte protein- og α-amylase-relaterede parametre. Således blev proteinindholdet øget med 0.4 til 1.7%-point, indholdet af vådgluten blev øget med 1.8 til 5.5%-point, Zeleny sedimentationsvoluminet blev øget med 1.4 til 3.5 mL, Farinograph vandabsorption blev øget med 0.5 til 1.4%-point og faldtallet blev øget med 10 til 48 s i melprøverne fra den bedste af tre fraktioner i forhold til udgangsmaterialet (Artikel II). Optimale prædiktionsmodeller af melets funktionalitet blev opnået ud fra nærinfrarød

refleksionsspektroskopi (NIR) på mel og partielle mindste kvadraters projektioner på latente strukturer (PLS) og multiblok-PLS (MBPLS). Nærinfrarød transmissionspektroskopi (NIT), infrarød (IR) refleksionsspektroskopi og nuklear-magnetisk resonans (NMR) relaksometri af bageprocessen og kombinationer af førnævnte teknikker gav anledning til modeller med større fejl. Melets funktionalitetsparametre var bedre end ovenstående spektroskopiske teknikker til prædiktion af brødets kvalitet, idet funktionalitetsparametrene indeholdt unik information som ikke blev modelleret tilstrækkeligt godt ved spektroskopi (Artikel III).

III. En ny kemometrisk metode, 2D PARAFAC-Laplace-opløsning, blev udviklet til unik bestemmelse, kvantificering og fortolkning af NMR-data fra 2-dimensionelle diffusions-relaksations-målinger af fedt- og vandfordeling og dynamik i dej (Artikel IV).

Table of contents

List of publications

Paper I
E. Tønning, L. Nørgaard, S.B. Engelsen, L. Pedersen, K.H. Esbensen (2006): Protein heterogeneity in wheat lots using single-seed NIT — A Theory of Sampling (TOS) breakdown of all sampling and analytical errors. *Chemometrics and Intelligent Laboratory Systems*, 84, 142-152.

Paper II
E. Tønning, A.K. Thybo, L. Pedersen, L. Munck, Å. Hansen, S.B. Engelsen, L. Nørgaard: Bulk quality diversification of organic wheat by single-kernel near-infrared (SKNIR) sorting. *Cereal Chemistry*, submitted.

Paper III
E. Tønning, A.K. Thybo, L. Pedersen, L. Munck, F. van den Berg, S.B. Engelsen, L. Nørgaard: Stepwise multivariate prediction of wheat flour functionality and bread quality. *Cereal Chemistry*, submitted.

Paper IV
E. Tønning, D. Polders, P.T. Callaghan, S.B. Engelsen (2007): A novel improved method for analysis of 2D diffusion-relaxation data – 2D PARAFAC-Laplace. *Journal of Magnetic Resonance*, 188, 10-23.

Additional publications
E. Dock, J. Christensen, M. Olsson, E. Tønning, T. Ruzgas, J. Emnéus (2005): Multivariate data analysis of dynamic amperometric biosensor responses from binary analyte mixtures - application of sensitivity correction algorithms. *Talanta*, 65, 298-305.

E. Tønning, S. Sapelnikova, J. Christensen, C. Carlsson, M. Winther-Nielsen, E. Dock, R. Solna, P. Skladal, L. Nørgaard, T. Ruzgas, J. Emnéus (2005): Chemometric exploration of an amperometric biosensor array for fast determination of wastewater quality. *Biosensors and Bioelectronics*, 21, 608-617.

List of abbreviations

AACC	American Association of Cereal Chemists
ANOVA	Analysis of Variance
ATR	Attenuated total reflectance
APLSR	ANOVA partial least squares regression
CGMP	Current Good Manufacturing Practice
CH_L	Constitutional heterogeneity
COST	Consider One Separate variable at a Time
CPMG	Carr-Purcell-Meiboom-Gill
DH_L	Distributional heterogeneity
DON	Deoxynivalenol
DRCOSY	Diffusion-Relaxation Correlation Spectroscopy
EMSC	Extended Multiplicative Signal Correction
FID	Free Induction Decay
FSE	Fundamental Sampling Error
GA	Genetic Algorithms
GSE	Grouping and Segregation Error
ICC	International Association for Cereal Science and Technology
ICS	International Chemometrics Society
iPLS	interval PLS
IR	Infrared
IUPAC	The International Union of Pure and Applied Chemistry
MBPLS	Multi-Block Partial Least Squares
MDA	Multiple discriminant analysis
MLR	Multiple Linear Regression
MSC	Multiplicative Signal Correction
NIPALS	Non-linear Iterative Partial Least Squares
NIR	Near-Infrared Reflectance (or Near-Infrared)
NIRS	Near-Infrared Spectroscopy
NIT	Near-Infrared Transmittance
NMR	Nuclear Magnetic Resonance
O/O2-PLS	Orthogonal Projections to Latent Structures
OSC	Orthogonal Signal Correction
PAC	Process Analytical Chemistry

PARAFAC	Parallel Factor Analysis
PAT	Process Analytical Technology
PC	Principal Component
PCA	Principal Component Analysis
PCR	Principal Component Regression
PGSTE	Pulsed Gradient Stimulated Echo
PLS	Partial Least Squares (regression) or Projection on Latent Structures
r	Correlation Coefficient
RMSECV	Root Mean Squared Error of Cross Validation
RMSEP	Root Mean Squared Error of Prediction
RF	Radio Frequency
RSD	Relative Standard Deviation
SKNIR	Single-Kernel Near-Infrared
S/N	Signal to noise (ratio)
SUO	Sampling Unit Operation
SVD	Singular Value Decomposition
T_1	Spin-lattice (longitudinal) relaxation time constant
T_2	Spin-spin (transverse) relaxation time constant
TAE	Total Analytical Error
TOS	Theory of Sampling
TPA	Texture Profile Analysis
USFDA	United States Food and Drug Administration

1. Introduction

Enjoying a soft and spongy slice of freshly baked leavened wheat bread is by many considered an exquisite gourmet experience as well as a fulfilling meal. Ever since the invention or discovery of yeast fermented dough some 5000 years ago by the Egyptians, wheat loaf bread has been a substantial part of the human diet – at least in some cultures. In the early days of bread making the raw material would be emmer (*Triticum dicoccum*) derived from the wild tetraploid, *Triticum dicoccoides*, as this was the most commonly cultivated wheat at the time. Emmer was later replaced by the hexaploid spelt (*Triticum spelta*) and another tetraploid; durum wheat (*Triticum durum*), which had higher yield and naked kernels, which made it easier to thresh. However, while durum wheats still has about 10% of the wheat market today, the rest is virtually covered by cultivars of hexaploid bread wheats, (*Triticum aestivum*)[1]. *T. aestivum* has the advantage of being highly adaptive and suited for a wide range of food products and feed (Belderok, 2000). It is cultivars of the latter which are the objects of observation in this thesis as they magically turn into dough and bread.

1.1. Diversification of food

Consumer demands are nourishing developments of a more diversified supply of high quality food products. Products with special attributes or characteristics are in high demand. Thus food products considered 'specialities' or with a special function, i.e. functional foods, can probably obtain higher prices in the marketplace, making products with a special story very lucrative for the industry. However, industrial processes have long been optimised for production of uniform products conforming to set specifications every day all year round, which certainly is a key quality attribute. Although consumers demand diversification and special treatment they also respond widely to recognition. E.g. a certain bread product should not change significantly in size, taste, texture or keeping qualities over the cause

[1] It should be mentioned that there is a growing commercial interest for the historic species such as einkorn (*Triticum monococcum*), emmer and spelt as bread ingredients due to their special characteristics in regards to functionality, taste and aroma (Larsen, 1999).

of the year due to variation in raw material or other production parameters. This type of random diversification would be detrimental for consumer confidence in the product as well as in the producing company.

1.2. Process analytical technology (PAT)

The two-fold demands to the food industry for both uniform and diversified quality products call for an increased focus of understanding the process from the raw material to the final product. That is the study of how raw material properties in combination with process control influence the properties of the final product. This area has always been a focus point in the industry in general due to bottom line considerations. Assessing raw material quality and innovative engineering in that respect is not a new phenomenon. The concept of process analytical chemistry (PAC) in which process monitoring and control of key manifest parameters are employed to predict and insure a certain end product quality during the process has been around for a century (Workman et al., 2005). However, it is the developments of new fast responding sensors, computers and advanced data analytical tools, which enable the industry and researchers to actually utilise the concept of PAC in full. At the turn of the century the United Stated Food and Drug Administration (USFDA) motivated a full integration of PAC in the pharmaceutical industry; "Process Analytical Technology (PAT) – A framework for innovative pharmaceutical development, manufacturing and quality assurance" (USFDA, 2004a). Furthermore PAT was introduced as a central part of current good manufacturing practices (CGMPs) in the industry (USFDA, 2004b). The PAT initiative (USFDA, 2005) was the first official regulatory acknowledgement of the potentials of the integrated technologies involved as the key strategy to ensure products of high quality and safety and is foreseen to be a turning point for all technologies involved and for all industries involving controlled physical, biological and chemical processes in the production chain – not limited to the pharmaceutical industry.

Although the historical definition of process analysis is limited to: "Chemical or physical analysis of materials in the process (stream) through the use of an in-line or on-line analyzer" or in short: "Analysis *in* the process", the more embracing PAT approach is encompassing all aspects of a process, in short: "Analysis *of* the process" (Guenard and Thurau, 2005, Workman et al., 2005) and considers a process a consecutive row of unit operation from raw materials to final products and wastes. Apart from the USFDA documents,

Bakaev (2005) has compiled a comprehensive introduction to the subject and Kourti (2006) reviews and emphasises the role of multivariate analysis in PAT.

1.3. PAT applied to the bread production process

Baking bread is such a process involving physical, biological and chemical unit operations and processes. The choice of seed grain determines the genetic background which fundamentally controls the outcome of the crop. Environmental conditions and the tending of the crop with water, fertilizer and pest control during growth is obviously important for nutrient uptake and quality development. Milling the grains to flour involves several steps and process decisions. The flour quality prior to the actual baking is thus a result of a series of events or unit operations and thus may vary a great deal from one lot to the other. Baking condition such as choice, quality and concentrations of the product ingredients, mixing time and intensity, rising time(s) and temperature(s), baking time and temperature are of great importance for the viscoelastic dough to form and develop during fermentation and baking into a tasty sponge of protein, starch, fat and water with a golden brown and crisp crust.

In order to gain process understanding and control which is the fundamental goal of PAT, relevant representative information must be gained by applying physical, chemical and biological analyses from the entire process from grain to the final perception of the consumer brain. By design, variability should be controlled and relevant information recorded. Manifest parameters, e.g. protein, gluten and falling number, and physicochemical fingerprints (spectra) recorded by spectroscopic multi-meters, such as near-infrared (NIR), infrared (IR) and nuclear magnetic resonance (NMR) are explored by multivariate data analysis and related to final product quality. This conglomerate of multifactor design, representative sampling, chemical, physical and biological analyses and multivariate "top down" analysis was established as a field of chemistry in its own right in the 1970's by Svante Wold and Bruce R. Kowalski (Kowalski, 1975). Chemometrics, which at the time was quite controversial with its 'shoot first – ask later' attitude to science or "analyse everything together, multivariately" as opposed to COST analysis (Consider One Separate variable at a Time) (Wold, 1991) is now an essential, integrated and evolving technology in the PAT framework (Kourti, 2006).

1.4. Aim

By emphasising the "process analytical perspective" in the title, this thesis focuses on the great potentials of utilising of chemometric technology (Munck, 2005) in a broad sense in the bread-making process. The subtitle; "Sampling, diversification, prediction and chemometric method development", pinpoint the different perspectives of the work covered by Papers I-IV.

1.5. Outline

Chapter 2 covers the process analytical toolbox i.e. three central PAT and chemometric elements utilised in this thesis; *representative sampling*, *measurements* of chemical, physical and biological parameters as well as multivariate physicochemical fingerprinting by spectroscopy and *data analytical methods*.

Chapter 3 covers the holistic view on the baking process including experimental design, the exploitation of the inherent heterogeneity of bulk wheat for quality diversification using single-kernel near-infrared (SKNIR) sorting, and the prediction of flour functionality and end product quality.

Chapter 4 explores the concept of NMR-baking by exploration of discrete and distributed exponential fitting for analysis of relaxing water and fat components in the baking process. An improved method, 2D PARAFAC-Laplace decomposition, for analysing 2-dimensional diffusion-relaxation NMR data is presented.

Chapter 5 concludes and put into perspective the results obtained regarding sampling, sorting, prediction, and chemometric method developments in the bread-making process and beyond.

2. The process analytical toolbox – chemometric technology

Process analytical technology concerns every aspect of a process (USFDA, 2004a). Chemometrics or chemometric technology (Munck, 2005) is a subset thereof – the toolbox for designing, sampling, analysing and understanding the process at hand. A handful of definitions of chemometrics presently co-exist without causing wide scientific turbulence. IUPAC states in The Gold Book: "Chemometrics is the application of statistics to the analysis of chemical data (from organic, analytical or medicinal chemistry) and design of chemical experiments and simulations" (McNaught and Wilkinson, 1997). The International Chemometrics Society (ICS) established in 1974 presently agree on this definition: "Chemometrics is the science of relating measurements made on a chemical system or process to the state of the system via application of mathematical or statistical methods." Svante Wold states that chemometrics is: "How to get chemically relevant information out of measured chemical data, how to represent and display this information, and how to get as much information into data" (Wold, 1995). The journal, *Chemometrics and Intelligent Laboratory Systems* (Chemolab), offers yet another definition: "Chemometrics is the chemical discipline that uses mathematical and statistical methods to design or select optimal procedures and experiments, and to provide maximum chemical information by analyzing chemical data." Whether the actual experiments including sampling and instrumentation is a part of these definitions or not is not completely clear – at least they are not identical. Petersen et al. (2005) however states that designing an optimal sampling plan in order to get representative samples can be viewed as a special case of the experimental design, hence making sure by design, that the samples provide maximum chemical information as opposed to unnecessary random noise and bias. Munck (2005) solved the instrumentation/measurement issue by broadening the concept as data does not exist without the actual physicochemical measurements (univariate as well as multivariate by spectroscopy or other multi-channel sensors) and named it: "Chemometric technology." This chapter on chemometric technology applied in the research presented in papers I-IV thus covers aspects of the The-

ory of Sampling (TOS), selected analytical methods and specific multivariate data analytical tools.

2.1. Democracy in science – representative sampling

Walking the corridors of virtually any scientific department around the world involved in recording or handling empirical data, the word 'sample' is often heard. Sometimes the sample is involved in very advanced state-of-the-art experiments being performed for the first time in the history. Sometimes the sample is part of a whole series of samples assembled to generate advanced calibration models in which future samples are to be estimated or evaluated. Sometimes the sample is the latest synthesis product of a chemical reaction. At other times the sample is a bag of material received by the test laboratory for analysis by some standard protocol. The term is treated and understood highly selective depending on scientific tradition and confusions often occur especially with respect to the quality and validity of the analytical results obtained by analysing those samples.

However throughout this text a sample is a specific part of something else – the lot[2]. Not just any part, but the part that ended up in the analytical volume based on democratic principles. A process in which all fragments, i.e. grains or flour particles had the same opportunity of ending up in 'parliament' being a physical average of the lot. In the sample cup parliament the representatively selected fragments are gathered to speak the case of the lot when 'interviewed' by either destructive or non-destructive analytical methods. The heterogeneity of the lot is by translation to the composite sample (Lamé et al., 2005) obeying the fundamental sampling principle (Gy, 1998) and ensuring physically unbiased samples for analysis.

Having a learned attitude towards samples and sampling, scientists and all other professionals working with empirical data develop a strategy for dealing with noise. The apparently random differences occurring when measuring the same material many times, either repeating the experiment on the same sample or by actually extracting new samples from the lot, are often

[2] The lot is all of the delimited material of which an average measurement is required. In Papers II and III for instance a lot is the entire collection of grains harvested in a specific growth year with a specific agronomical treatment or the entire collection of grains with a defined quality fractioned by the TriQ sorter.

confused in terms like analytical errors, measurement errors or sampling errors. The uncertainties accompanying all empirical results are thus not always understood. Specific results may sometimes be accompanied by the uncertainty supplied by the instrument manufacturer or by calculating the variance from replicate experiments or measurements. This could be both right and wrong depending on what is reported and why.

Grabbing a sample from a lot or process of particular interest and bringing it to either in-house or external testing, measurements or recording involves costs. Some tests may even be very costly or crucial for the parties involved, hence spending just a little extra time on ensuring the quality of the sample may be worthwhile (Gy, 1986). Gy (1995b) hit the nail on the head by citing Kaye (1967): "The accuracy of many analytical data reports is a mirage because unwitting negligence and false cost consciousness have ensured that a sample of powder taken with cursory swiftness has been examined with costly precision"

2.1.1. The Theory of Sampling (TOS)

The key to understand empirical measurements, be it univariate as protein content of wheat or multivariate as a near-infrared spectrum, is to realise that "analytical results are *estimates* of unknown quantities" (Gy, 1995a) derived in a multi-step process of sampling ending up with what we understand as the analysis:

Estimation = sampling + analysis (Eq. 2.1)

The theory of sampling (Gy, 1986, 1995a, 1995b, 1998, 2004a, 2004b and 2004c, Heydorn and Hansen, 2005, Pitard, 1993) can be viewed as a special case of probability and population statistics in which each fragment or otherwise delimited part in the lot is considered an experimental unit. Even though Pierre Gy and his co-workers developed TOS more or less alone and against all odds (Gy, 1995a, 2004d) over a number of decades to overcome practical and theoretical obstacles for obtaining truly representative samples from more or less complicated materials, it is in no way in conflict with statistical theory. Lwin et al. (1998) cite others such as Wilson (1964) and Ingamells and Switzer (1973) for their independent discovery of the theory, however Gy is considered the most dedicated to establish a comprehensive and accessible theory applicable for analysts at all levels. TOS can be consid-

ered applied statistics both in terms of understanding the theoretical prerequisites of obtaining a representative sample, designing the sampling tools and right down to evaluating the quality of the estimates by assessing the variances and biases of the results.

2.1.2. Accuracy, reproducibility and representativity

Sampling is a mass reduction process in which a small part of the lot is selected and subjected to analysis, with an objective goal of obtaining an accurate and reproducible estimate of the true average value of the entire lot. If the sampling process produces results which are systematically deviating from the true value, i.e. either too high or too low, the difference is termed *bias*; hence an analytical result with high bias has low accuracy. The random deviation from the mean analytical result is assessed by the standard deviation; hence a sampling process producing analytical results with a relatively large standard deviation has a low reproducibility. Gy (2004a) defined the sum of the squared bias and squared standard deviation as an objective measure for representativity. Thus optimising the sampling process with respect to both accuracy and reproducibility leads to more representative results. To answer the question whether an analytical result is representative or not, the analyst has to define acceptable levels of accuracy, reproducibility and representativity. The formal theory and relevant equations are presented and applied in Paper I.

2.1.3. Properties of composite fragmented materials

Wheat grain lots and wheat flour lots are materials composed of fragments of varying size, shape, surface properties, density and composition. Such lots are zero-dimensional (0D) as all fragments (grains or flour particles) are equally accessible from a sampling point of view, although the assembly of fragments in a stock pile or bag indeed appears three-dimensional to the analyst. A multi-phase continuous material such as dough or bread crumb may also be considered 0D as long as all parts of the material are equally accessible, hence conforming to the fundamental sampling principle. Further on lot dimensionality (0D, 1D, 2D and 3D) can be found in Gy (1998) and Petersen et al. (2005). Cases of process sampling and 1D sampling in relation to PAT can be found in Esbensen et al. (2007), Holm-Nielsen et al. (2006) and Petersen and Esbensen (2005).

The challenge in regards to extracting a truly representative sample relates to the uneven distribution of the analyte in the lot, e.g. the protein content is not the same in each individual wheat kernel as demonstrated in Paper I and utilised in Paper II and III. This fundamental variation between individual fragments defined as the constitutional heterogeneity of the lot (CH_L) is the origin of the fundamental sampling error (FSE) and is responsible for the theoretically lower limit of the sampling reproducibility. An uneven distribution of particles gives rise to the distributional heterogeneity of the lot (DH_L). This distribution is caused by several factors, e.g. when a field of ripe wheat is harvested, the quality variations in the field are to a large extent preserved in the storage due to the autocorrelation of adjacent grains. Neighbouring grains simply tend to accompany each other hence preserving the distributional heterogeneity from the field during harvest and storage. In addition segregation of small, dense and slippery particles as opposed to large, light and rough particles during transport and storage leads to a more or less segregated lots and thus adds to the distributional heterogeneity of the lot. DH_L is the origin of the grouping and segregation error (GSE) and either adds to or subtracts from the reproducibility measure[3]. DH_L is also the cause to bias, when samples are extracted without proper attention to the sampling process, e.g. by grabbing the samples at easy accessible sites rather than ensuring a truly representative sampling process.

Paper I is a total breakdown and quantification of the sampling and analytical errors in bulk wheat sampling and serves as a reference study in cereal sampling and emphasises the contributions to the order of magnitude of FSE and GSE by quantification of variances and biases of both correct representative sampling and incorrect grab sampling.

2.1.4. Seven sampling unit operations

The theory of sampling prescribes seven sampling unit operations (SUOs) which enable the enlightened analyst to estimate properties of any given lot and by reference to the sampling process substantiate that the results obtained are both unbiased and reproducible. The seven SUOs are presented

[3] The distributional heterogeneity may cause grouping of uniform material in which grab sampling may result in lower variance of replicate samples than would be obtained by correct representative sampling, hence leading to a too optimistic reproducibility measure and an unknown bias.

below and not necessarily in the order given elsewhere (Petersen et al., 2005), but rather in the order of importance or practical implementation followed by a short guidance.

SUO 1: *Always obey the fundamental sampling principle*. Take a careful overview of how you are actually going to achieve the analytical volume from the entire lot. Does this process honestly ensure that all fragments have the same probability of ending up in the sample? By doing so the bias is reduced to zero and sleepless nights are avoided.

SUO 2: *Mix the lot*. Mixing the lot prior to the mass reduction reduce the autocorrelation of fragments, hence reduce the GSE.

SUO 3: *Use composite sampling*. This can be done easily by utilising commercially available dividers or splitters. These devices ensure accuracy and reproducibility with a minimum of labour and should always be preferred for laboratory scale mass reduction (Petersen et al., 2004). When going from stockpile to laboratory scale (a few kg) always go for a large number of increments when composing the sample. Don't be mean with size if not constrained. Composite sampling may also be achieved by averaging measurements of several increments, as done in Paper II and III when texture profile analysis is performed on several slices of the bread crumb. Increasing the number of increments will reduce the GSE.

SUO 4: *Comminute whenever necessary*. Comminution at selected stages of the mass reduction process is beneficial for reducing the FSE and simultaneously the GSE. The FSE and GSE of a flour sample are much smaller than the corresponding grain sample of the same analytical mass.

SOU 5: *Perform a heterogeneity characterisation for new materials and new sampling processes*. Faithfully repeat the entire sampling process several times (at least ten times) in order to achieve the variance of the analytical result, i.e. the variance of the global estimation error. That is the reproducibility to the order of magnitude. More thorough heterogeneity characterisations may be necessary for trouble shooting the process for sampling steps generating unnecessary large errors.

SOU 6: *Turn large 2D and 3D lots into 1D for better access*[4]. Even better just perform incremental sampling (SOU 3) while the lot is laid up in the first place.

SOU 7: *Perform variographic characterisation of 1D heterogeneity*. In processes variography may elucidate systematic temporal and spatial variations of great importance (Esbensen et al., 2007, Petersen and Esbensen, 2005).

Not all operations may be needed; rather a selected number of operations are needed depending on the problem at hand. An illustrative example of a heterogeneity characterisation was given by Whitaker et al. (2000) by determining the error contributions in the determination of deoxynivalenol (DON) in wheat using representative sampling by splitting and comminution in the effort to reduce masses in 20 kg lots to 25 g analytical volumes. The tested protocol showed that 22 % of the total variance, i.e. the variance ratio, was due to primary mass reduction by representative splitting 20 kg down to 454 g. The variance ratio of grinding the 454 g and automatically reduce mass to 25 g was 56% while the variance ratio of the analysis was 22%. The total coefficient of variance (CV) of the global estimation error, i.e. of the total variance, was 13.6% for a 5 ppm sample, which was considered relatively low for this type of measurements. However attention to the comminution and automatic sampling mill may improve the global estimation error. If the device is truly providing representative samples, an idea was to increase the analytical volume as suggested by Whitaker et al. (2000). Another multi-stage sampling process in particulate materials involving several crushing and mass reduction operations was presented by Lwin et al. (1998). In environmental research, studies have shown that contaminants can be extremely unevenly distributed. Lamé et al. (2005) showed that by taking a small sample for heavy metal analysis without proper attention to heterogeneity there is only a slight possibility, that the obtained analytical result will be a good estimate for the mean concentration. Hence proper

[4] The dimensionality of the lot has not been covered here as lots regarded in this work are all considered zero-dimensional and thus fully accessible. Sometimes 0D-lots are so big, e.g. a stock pile, that turning them into 1D by moving it on a conveyor belt is a practical and pragmatic solution to make all parts of he lot equally accessible.

sampling protocols are finding their ways into legislation to avoid the above mentioned "casino effect" - at least in The Netherlands.

Mixing (SUO 2) is often sought of as an adequate operation for stating subsequently grabbed samples to be representative. However, although mixing is disturbing the autocorrelation present, it does not remove autocorrelation. Mixing rather creates a new state of autocorrelation which usually has a lower distributional heterogeneity. One must keep in mind that the mixing operation used may have the direct opposite effect causing increased grouping or segregation. Thus stating a sample grabbed from a well mixed lot as representative could be a self-deception if not carefully substantiated.

The main problem in assessing the true bias, i.e. the systematic deviation from the true average, is that the true value is unknown – and in most cases will remain so. That is probably why SUO 1 is so often neglected. However, model experiments with known quantities of constituents that do not react with each other clearly show that obeying the fundamental sampling principle is worthwhile. Mass reduction devices and protocols can in such cases be evaluated with respect to representativity, both in terms of reproducibility and accuracy. Such studies show that true splitting devises such as riffle splitters, rotational dividers and similar composite sampling strategies outperform all other methods for establishing unbiased analytical masses for analysis. Grab sampling compromise SUO 1 and should at all costs be avoided as they invariably produce biased results (Allen and Khan, 1970, Gerlach et al. 2002, 2003, Petersen et al., 2004, Smith, 2004, Venables and Wells, 2002).

2.1.5. Heterogeneity equals diversity

Although the heterogeneity observed in cereals and in all other parts of the physical world causes some theoretical and practical challenges for obtaining average properties, heterogeneity or diversity opens up new possibilities. In Paper II and III the heterogeneity of apparently homogeneous wheat lots are exploited by utilising dedicated equipment, the TriQ sorter, for measuring and sorting single fragments, i.e. wheat grains, according to an internal complex quality trait of each individual kernel (Dowell et al., 2006a, Löfqvist and Nielsen, 2003, 2004, Munck, 2008, Nielsen, 2002). This technology opens up great perspectives for advanced 'homogenisation' and im-

provements for all sorts of composite and particulate materials by sorting (Chap. 3).

2.2. Physicochemical standard analyses of wheat

A great number of standard analyses are endorsed by the American Association of Cereal Chemists (AACC) and the International Association for Cereal Science and Technology (ICC) for characterisation of wheat and other cereals in a uniform and comparable way throughout the world. A few of these were utilised in Paper II and III to characterise the flour functionality. The chosen parameters cover central chemical, physical and biological properties; however do not encompass a complete characterisation, which was not possible according to the available nails principle at the time of analysis.

2.2.1. Protein, moisture and ash

Fast determination of flour protein content, moisture and ash is nowadays routinely performed using indirect measurement. Near-infrared transmission (NIT) spectra are recorded and the protein content calculated according to a global multivariate calibration in the Foss Infratec™ 1241 Grain Analyser fitted with a flower cup module. The principle recommended by ICC (ICC Recommendation No. 202) is based on multivariate calibration to the original determinations of protein using the Kjeldahl method, oven drying and combustion. The AACC has several approved methods based on near-infrared analysis.

2.2.2. Wet gluten

Wet gluten content and gluten index are determined by washing the flour during mixing with subsequent centrifugation according to AACC Standard No. 38-12 and ICC Standard No. 155. Gluten consists mainly of protein (90%), lipids (8%) and carbohydrates (2%), the latter mainly water-insoluble pentosans capable of binding large amounts of water (Belitz and Grosch, 1999). The gluten is responsible for the viscoelastic properties of wheat dough and forms the air-holding polymeric network which is fixed into a spongy crumb structure during baking.

2.2.3. Zeleny sedimentation volume

The sedimentation volume according to Zeleny, AACC Method 56-61A and ICC method 116/1, is a fast obtainable index for dough-mixing characteristics and baking quality (Pinckney et al., 1957) of flour on a 14% moisture

base. The swelling volume of the gluten fraction in a diluted solution of lactic acid and alcohol is measured in [mL]. Both high gluten content and good gluten quality result in high sedimentation values and may vary from 8 for low and weak gluten content to 78 for high and strong gluten content.

2.2.4. Farinograph

The rheology and water absorption of flour may be assessed by a dedicated mixing device called a Farinograph. Farinograph water absorption, development time, stability and softening according to ICC standard no. 115/1 and AACC method no. 54-21 are measures of flour water absorption to a specified consistency and the mixing properties and tolerance of the resulting dough. The measurements are carried out on a Brabender Farinograph instrument with a specially designed, thermostatic controlled mixing chamber for either 300 g or 50 g of flour on a 14% moisture base. A time-consistency curve is generated with a characteristic shape reflecting the internal quantity and quality of protein and starch. In Fig. 2.1, farinograms of bisquit, feed, hard bread and commercial bread wheat are shown. The consistency is measured in arbitrary Farinograph units, FU, and water addition is adjusted so that the maximum consistency is 500 FU. The added water corrected for moisture content of the flour is the water absorption. Development time, stability and softening are read on the farinogram according to Fig. 2.1A and represent the quality of the viscoelastic gluten network formed during mixing.

The Farinograph water absorption is positively correlated to both protein content and the amount of damaged starch (Delwiche and Weaver, 1994, Mirablés, 2004). Intact starch granules absorb generally one third of their weight in water, while damaged starch may increase their weight up to three times (Mirablés, 2004). The amount of damaged starch is determined by the milling process and the grain hardness. Starch granules in hard wheat grains tend to fracture easer thus leading to a higher starch damage rate. Since damaged starch provide extra sugar to the fermentation process through easier degradation by α-amylases, increased bread volume is expected, however too much starch damage leads to small bread volume with a heavy crumb structure. A balance is preferred. Flours with strong gluten usually have a long development time and a high mixing tolerance reflected in a large stability and a low softening.

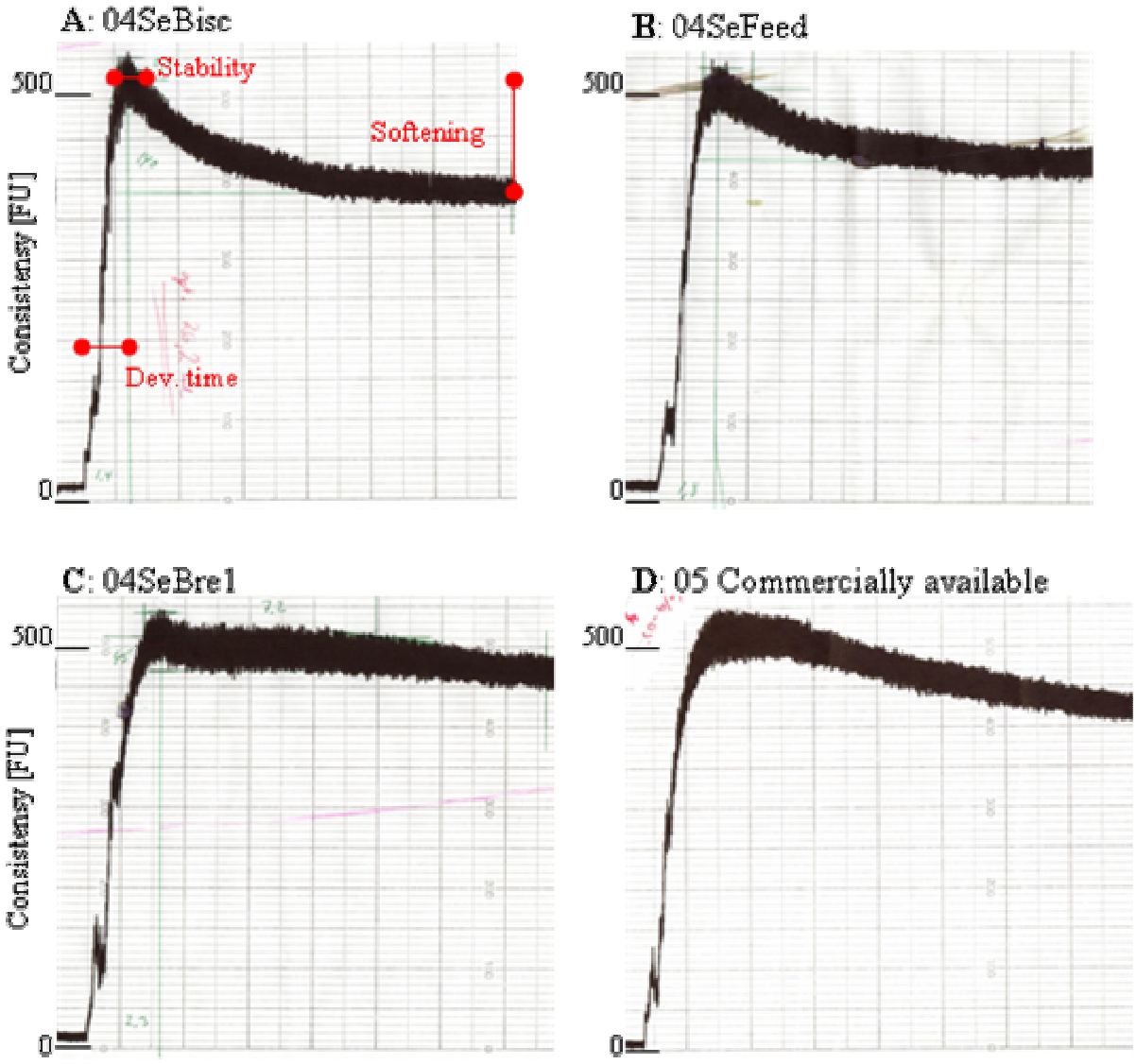


Figure 2.1: Typical farinograms with increasing development time. A: Relatively soft biscuit wheat (04SeBisc) with short dev. time, low stability, high softening. B: Fodder wheat (04SeFeed) with short dev. time, low stability and some softening. C: Hard bread wheat (04SeBre1) with intermediate dev. time, high stability and low softening. D: Commercially available bread wheat with long development time, intermediate stability and intermediate softening.

2.2.5. Amylograph

Amylograph gelatinisation maximum temperature, maximum viscosity and beginning gelatinisation temperature according to ICC standard no. 126/1 and AACC method no. 22-10 are measures of the starch gelatinisation properties and simultaneously the α-amylase activity on a 14% moisture base. A water suspension of flour is stirred in the Brabender Amylograph rotating bowl while the temperature is increased linearly at constant heating rate of 1.5°C/min. The heating rate corresponds to the heating rate during bread

baking and is thought to provide realistic information of the gelatinisation during bread baking. The viscosity is measured continuously and the peak viscosity in arbitrary Amylograph Units (AU) and the corresponding temperature is read on the amylogram. High beginning temperature results from low amount of damaged starch, while high maximum gelatinisation temperature and high maximum gelatinisation is a result of low α-amylase activity. Compared to a real baking process the Amylograph suspension experiment is working at excess water which is not the case in bread were starch granules are only partly wetted due to water deficit (Hardacre, 2006).

2.2.6. Falling number

Falling number determination according to ICC standard no. 107/1 is a fast method for determining α-amylase activity in flour on a 14% moisture base. After a fast gelatinisation of the starch in an aqueous suspension of flour in a boiling water bath, the time consumption in seconds for a stirrer to fall through the gel undergoing liquefaction is determined. High activity results in low falling number. The falling number analysis is a fast and simple counterpart to the more laborious Amylograph measurement.

2.2.7. The 14% moisture base

The often applied 14% moisture base is a standard base for many flour tests probably developed for practical and traditional reasons. Flour always contains a certain amount of water typically in the order of 10-16% as it is hygroscopic by nature. Using the 14% moisture base practically means that the flour amounts used, $m_{F(14\%)}$, in the tests above has to be corrected for actual moisture content, $m_{F(\text{Actual moisture})}$, in order to work with a standardised amount of dry matter:

$$m_{\mathrm{F(Actual\ moisture\%)}} = m_{\mathrm{F(14\%)}} \cdot \frac{100\% - 14\%}{100\% - \mathrm{Actual\ moisture\%}} \qquad \text{(Eq. 2.2)}$$

In tests involving water addition, such as the Farinograph water absorption, the actual water content has to be taken into account when calculating water absorption on the basis of added water, m_W, resulting in maximum consistency of 500 FU:

$$\mathrm{Water\ absorption}(14\%) = \frac{m_{\mathrm{F(Actual\ moisture\%)}} + m_{\mathrm{W}} - m_{\mathrm{F(14\%)}}}{m_{\mathrm{F(14\%)}}} \cdot 100\% \qquad \text{(Eq. 2.3)}$$

2.3. Texture profile analysis (TPA)

Objective determinations of food texture or mouth feel has long been a focus area for food technologist and the food industry, as texture is a fundamental quality attribute of food which determines the consumer acceptability to old and new products. Szczesniak (1963b) reviews a number of fundamental, empirical and imitative instruments of which many are still in use. The texture profile analysis (TPA) principle was introduced by Friedman et al. in 1963 and has become a standard procedure for objective sensory evaluation of texture of various food products (Bourne, 2002, Bourne and Comstock, 1981). It has many advantages to sensory panel evaluation. Experiments are low cost and easily standardised and reproduced. In addition it shows good correlations to sensory attributes (Friedman et al., 1963, Henry et al., 1971, Szczesniak, 1963a, Szczesniak, 1968, Szczesniak et al., 1963). This was also observed in the present study (Chap 3.3, Papers II and III).

The analysis imitates the chewing mechanism of humans, by compressing and decompressing a standardised food sample with a piston moving at constant speed down and up twice with a pause in between. It combines as such the empirical and imitative principle in a relatively simple way (Szczesniak, 1963b). The force used to maintain constant speed is recorded continuously throughout the 'chewing' time. A time-force curve appears and traditionally specific features are extracted and calculated in order to characterise the sample. The mean time-force curve of the bread crump using a 40 mm diameter cylindrical steel probe with sharp edges on four slices of bread sample, (04SeBre1) is shown in Fig 2.2.

2.3.1. Feature extraction

The texture features which are typically extracted from the profiles like the one in Fig. 2.2 are listed in Table 2.1. Hardness 1 and 2 are forces used at maximum compression in first and second bite respectively. Resilience is the ratio between the work done from anchor 2 to 3 under decompression and the work done compressing the sample in the first place from anchor 1 to 2. If a sample bounces right back up in the same way it was compressed the resilience would be close to 1. Cohesiveness is much like resilience but based on the ratio between the second down stroke work and the first work, i.e. the areas under the down stroke curves, leaving a little time, typically 5 s, for the

sample to recover before compressing the sample the second time. The springiness describes the recovery height relative to full compression height. If compressed only very little, say below 50%, fresh bread can have resilience, cohesiveness and springiness close to 1, as the crumb structure stays intact. These parameters are highly correlated in bread crumb, as they are featuring more or less the same physical phenomenon of recovery of a spongy product (Paper II).

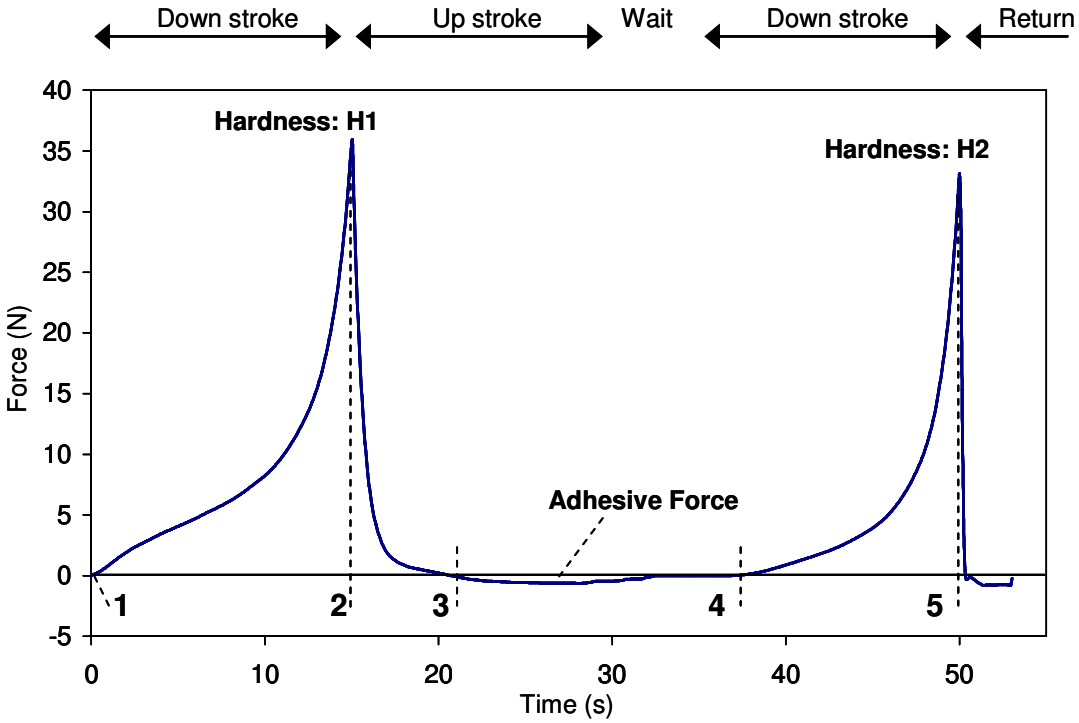


Figure 2.2: Average time-force curve of a texture profile analysis (TPA) of four bread slices. Three force parameters, Hardness: H1 and H2 and adhesive force (AF) are read directly at maximum compressions of 1st and 2nd 'bite' and at maximum negative force during the 1st up stroke. Anchors 1 to 5 are set to determine areas and lengths for calculation of further texture parameters (Table 2.1).

Gumminess and chewiness are interaction terms involving hardness 1 and may be considered as an arbitrary measure for the energy needed to disintegrate a solid and masticate a semisolid food respectively (Bourne, 2002).

Adhesive force is the maximum negative force of the bread sticking to the base plate and the moving piston on its way up from the first compression. Very sticky samples will exhibit large adhesive force. Adhesiveness is a similar feature, just based on the entire negative work exhibited, i.e. the area over the curve from anchor 3 to 4. The sample in Fig. 2.2 was the stickiest

sample in the entire investigation and thus illustrates the phenomenon even though wheat bread is usually not very sticky. Another parameter not included in this study, stringiness, is used in semisolid foods such as ketchup, cream cheese and pudding and refers to the distance the sample is extended under decompression from anchor 3 in Fig 2.2 until breaking off (Henry and Katz, 1969). Other parameters regarding the negative part of the time-force curve can also be derived (Henry et al., 1971). In some harder foods such as apples, biscuits and potatoes an additional peak appears during the first down stroke due to fracturability (Bourne, 2002).

Table 2.1: Some bread crumb texture profile parameters derived from the time-force curve of TPA. Numbers refer to the anchors in Fig. 2.2.

Texture paramters:	Abreviations and calc.	Units
Hardness 1	H1	N
Hardness 2	H2	N
Adhesive Force	AF	N
Resilience	Re = Area23/Area12	-
Cohesiveness	Co = Area45/Area12	-
Springiness	Sp = Length45/Length12	-
Gumminess	Gu = Co·H1	N
Chewiness	Ch = Gu·Sp	N
Adhesivenes	Ad = Area34	Ns

2.3.2. Compression rate

How much to compress the food is critical as it affects the assessment of the analysis. However, degrees of compression of 10 to 90 % have been reported and the effect of various compressions has been investigated (Bourne and Comstock, 1981). The more the food sample is compressed, the more information is gained in terms of a more detailed curve shape. Most parameters are dependent of the degree of compression, with fracturability as the general exception. Springiness and cohesiveness for some commodities appears to be less sensitive to degree of compression (Bourne and Comstock, 1981).

The upper limit of relative compression varies with food properties, but getting too close to this limit should be avoided. The recording of very hard foods may be halted if the instrument is overloaded and even small variations in the sample thickness may result in large variations in the recorded texture profile.

2.3.3. Correction for varying sample height

Recordings of the four slices of bread sample 03AaViWV are superimposed in Fig. 2.3 to illustrate the effect of varying slice thicknesses. The available slicing equipment for this study, a meat slicer with a rotating blade, had only a limited reproducibility, thus variation in slice thicknesses were inevitable. The thicknesses represented in Fig. 2.3 are 17.8 mm, 20.3 mm, 20.7 mm, and 23.0 mm which represents the upper limit of thickness variation found in this work. When compressed by 15 mm the degree of compression are 83.8%, 73.9%, 72.5% and 65.2% respectively. Clearly the thin slice was too thin as the recording stopped due to system overload. The successful measurements however still suffer from variation in actual compression hence the hardness parameters will be ill determined based on these data. An alternative compensation for this was introduced in the following way. The points at which the slices were actually compressed by 60% were found and the force extracted, $F_{60\%}$. These points are marked by stars in Fig. 2.3 and are between 10 and 15 s along the time axis depending on slice thickness. $F_{60\%}$ however is derived from samples which are still of varying thickness. 60% compression of 23.0 mm is 13.8 mm, while 60% of 17.9 mm is only base on a down stroke of 10.7 mm. Thus the force expected to be used on the individual slices if they would have been 20 mm, $F_{60\%,20\mathrm{mm}}$, can be pragmatically estimated from the slice thickness d:

$$F_{60\%,20\mathrm{mm}} = \frac{F_{60\%}\,20\mathrm{mm}}{d} \qquad \text{(Eq. 2.4)}$$

Thus a linear relation ship is expected between thickness and force used to compress a given sample 60%. This was not explored, but reasonably substantiated by improved mean relative standard deviations of determinations of $F_{60\%,20\mathrm{mm}}$, RSD($H1_{60\%,20\mathrm{mm}}$) = 21.7% and RSD($H2_{60\%,20\mathrm{mm}}$) = 21.0% as compared to $F_{60\%}$. RSD($H1_{60\%}$) = 22.2% and RSD($H2_{60\%}$) = 23.7%. This is indeed below RSD of the original hardness determinations of RSD(H1) = 40% and RSD(H2) = 39%. Thus H1 and H2 in the following imply $H1_{60\%,20\mathrm{mm}}$ and $H2_{60\%,20\mathrm{mm}}$ when recorded on the TA-XT2 Texture Analyser.

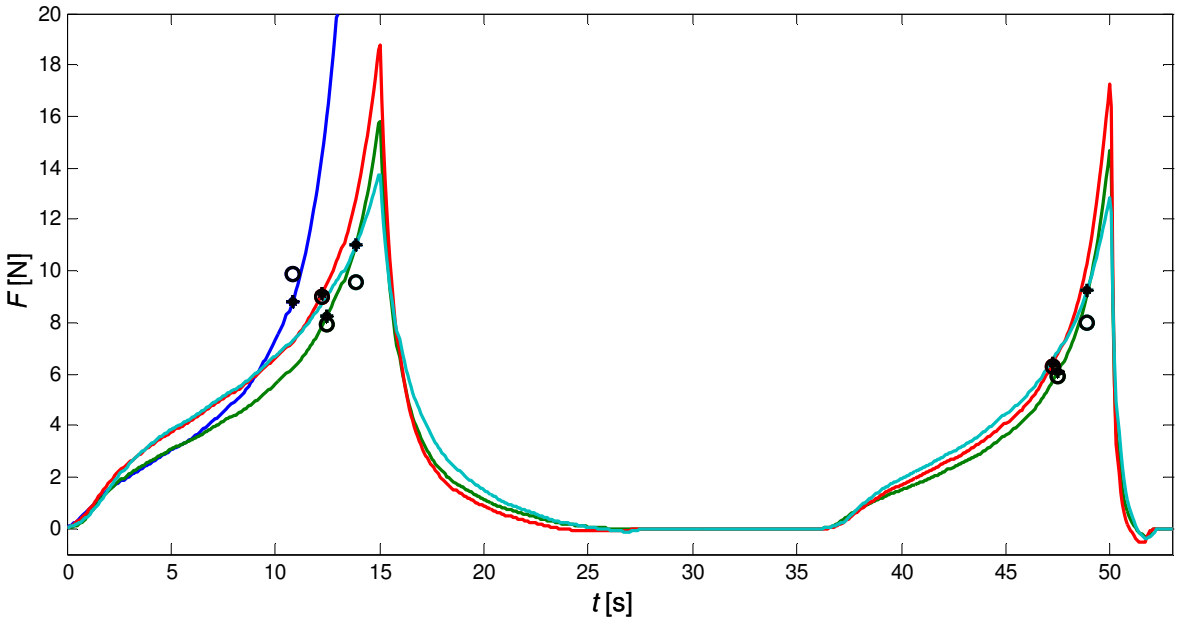


Figure 2.3: Time-force curves of four bread slices of bread sample 03AaViWV with varying thicknesses, — 17.9 mm, — 20.3 mm, — 20.7 mm, — 23.0 mm. The stars (*) are at 60% compression with respect to slice thicknesses. The circles (○) mark the vertical displacement when 60% compressions are normalised with slice thicknesses.

2.4. Sensory texture profiling

The ultimate quality test of food is sensory profiling by a panel of trained assessors. The 'instrument' does not differ much from other instruments in that a sensory panel return signals on request when presented to a sample. The sensors are human and the responses are given on an arbitrary scale depending on the attribute and setup, typically from 0 to 15. Thus the sensory panel must be trained (calibrated) to respond consistently and uniformly to the samples. The training involves both a conceptual consensus on the attributes to be measures and an appropriate use of scale for the test at hand. Sensory profiling is thus an objective assessment of food quality as opposed to preference studies. From a commercial point of view, preference tests are sometimes considered the ultimate tests of foods – and may be used in combination with sensory profiling to pinpoint key attributes decisive for consumer preferences.

Sensory profiling involves all five senses; sight, touch, smell, taste and hearing, and relies on the brain to translate and separate the product stimuli into sensations, perceptions and responses on an arbitrary scale from none (0) to high sensation (15). Texture is determined by the touch of the product, either by the mouth or fingers or a combination and can be divided into kinaes-

thetic and tactile senses (Meilgaard et al., 1991). In this thesis six texture properties of the bread crump were assessed. The mechanical properties, elasticity and fracturability were assessed by the fingers while hardness, chewiness and adhesiveness were assed by the mouth – all assessed by kinaesthetic senses. The moisture property on the other hand assessed in the mouth as dryness was assessed by tactile senses. Four additional properties was also assessed by the panel; yellowness (colour) of the bread crumb, aftertaste and cereal aroma and taste. The last two however were excluded in the investigations of Paper II and III due to low discriminative power between the samples. A description of the sensory attributes can be found in Paper II.

However well the sensory panel is trained, it is not possible to calibrate an entire panel to respond exactly uniformly with respect to all attributes. Different assessors thus utilise the scale differently with respect to level and range for each attribute. Data from sensory profiling thus need special attention in pre-processing and selection in order to extract all relevant information and suppress noisy and irrelevant information. The sensory data in Paper II and III has been evaluated using Martens et al. (2000) guide (Martens et al., no year) to remove level and range effects as well as assessors and attributes with no discriminative power. An ANOVA partial least squares regression (APLSR) was calculated using design variables (0/1) for assessors and replicates as **X** and the sensory variables as **Y**. Thus **X** consisted of eleven variables corresponding to ten assessors and one replicate variable. The APLSR models the part of **Y** which is associated with assessor and replicate level effects while the residual **Y** variation contains structure as well as noise concerning the samples. The operation is equivalent to mean centring every individual attribute over samples within each judge and over replicates across all judges and samples. The **Y** residuals are subsequently used as level corrected sensory data for further analysis and evaluation of judges and attributes discriminative powers (Martens et al., 2000). Elaborations on individual assessor and attribute performance assessed by signal to noise ratios can be found in Appendix A. The level corrected data were averaged over judges and replicates; hence scales are not the original in Paper II and III.

2.5. Spectroscopy

According to IUPAC in The Gold Book, spectroscopy is: "The study of physical systems by the electromagnetic radiation with which they interact or that they produce. Spectrometry is the measurement of such radiations as a means of obtaining information about the systems and their components. In certain types of optical spectroscopy, the radiation originates from an external source and is modified by the system, whereas in other types, the radiation originates within the system itself" (McNaught and Wilkinson, 1997).

Electromagnetic radiation interacts with physical systems in various ways depending on the energy of the radiation and the available excitation states present in the system. The great advantages of spectroscopy in science as well as in industry are already well established and especially in combination with multivariate data analysis (Bakeev, 2005, Bro et al., 2002). Especially low energy techniques are of great advantage as they are non-destructive and intact food samples can be characterised in a split second. Infrared (IR) and near-infrared (NIR) radiation interact with the vibrational states of covalent bonds as well as molecular rotational states (Osborne et al., 1993). Low energy radio waves interact with spin states of nuclei in the matrix which may be observed with nuclear magnetic resonance (NMR) spectroscopy (Callaghan, 1991).

2.5.1. Vibrational spectroscopy

Vibrational spectroscopy is of particular interest for research and industrial applications of various products since covalent bonds between atoms H, C, N, O and S, the constituents of all organic material, absorb light in the near-infrared (NIR) and infrared (IR) region of the electromagnetic spectrum to change only the vibrational states of the molecules. Herschel (1800) was the first to discover the invisible "heat" radiation of the sun which we refer to as near-infrared light. Now some 200 years after, NIR, IR and Raman spectroscopy is used in process monitoring and control (Bakeev, 2005) as well quality analysis of various products and raw materials. The spectrometers available are very accurate and reproducible and deliver physicochemical fingerprints packed with information which can be readily extracted by chemometrics (Munck, 2005).

In the IR region ranging from 4000 to 400 cm^{-1} (or 2500 nm to 25 μm), often referred to as the mid-IR, absorption changes the vibrational state of a molecular bond from the ground state to a discrete higher energy level (Coates, 2005, Williams and Flemming, 1995). The absorption frequency is determined by the masses of the involved atom as well as a number of other factors, such as the type of vibration (symmetric stretching, asymmetric stretching, bending, rocking, wagging and twisting), other groups attached to the atoms, temperature and pH. Different functional groups have distinct absorption bands; however the entire molecule has a distinct vibration or fingerprint.

While IR spectra in principle contain all bands of the fundamental vibrations[5], the near-infrared spectra contain only overtone and combination bands of the most anharmonic vibrations, which primarily mean bonds and groups with hydrogen (C-H, O-H, N-H and S-H). Only pure overtones of the stretching vibrations are seen in the NIR region, the remaining vibration types are only represented as combination tones in the NIR spectra. The holographic overtone pattern of especially the stretching vibrations only overlaid by combination tones was nicely demonstrated by Pedersen and Engelsen (2001) for ethanol and is also observable from Fig. 2.4A and B in the NIR and IR spectra of wheat – although immensely more complicated than ethanol. The bands at 1465 and 1000 nm correspond to the 1st and 2nd overtones of the fundamental overlapping stretching vibration of O-H and N-H in the IR spectrum around 3300 cm^{-1} (Peaks **1**), while the bands around 1780, 1200 and 920 nm corresponds primarily to 1st, 2nd and 3rd overtones of the fundamental stretching vibration of C-H in the IR spectrum around 2927 cm^{-1} (Peaks **2**). This repeating pattern was suggested to be part of a harmonic series by Coblentz (1905), hence the concept of overtones. The combination tones from 1900 to 2500 nm completely dominate the NIR spectrum (Fig. 2.4A) with dominating bands of O-H and N-H combinations at 1934 nm (Peak **6**), amid combinations around 2100 nm (Peak **7**) and C-H stretching combinations from 2280 to 2330 nm (Peaks **8**). In Paper III the NIR and IR spectra are explored for prediction of flour and bread quality. However, only the fingerprint region of IR from 1900 to 700 cm^{-1} was utilised (Fig.

[5] The more rigid bonds, symmetric bond and bonds with low dipole moment are usually not very intense in IR, but may be visible in the complementary Raman spectra (Jestel, 2005, Williams and Flemming, 1995).

2.4B). This part contain the Amid I and II bands (Peaks **3** and **4**) and characteristic C-O and C-N stretching bands (Peak **5**). A comprehensive assignment table of the IR region can be found in Williams and Flemming (1995), while Siesler et al. (2002) and Osborne et al. (1993) cover the NIR region.

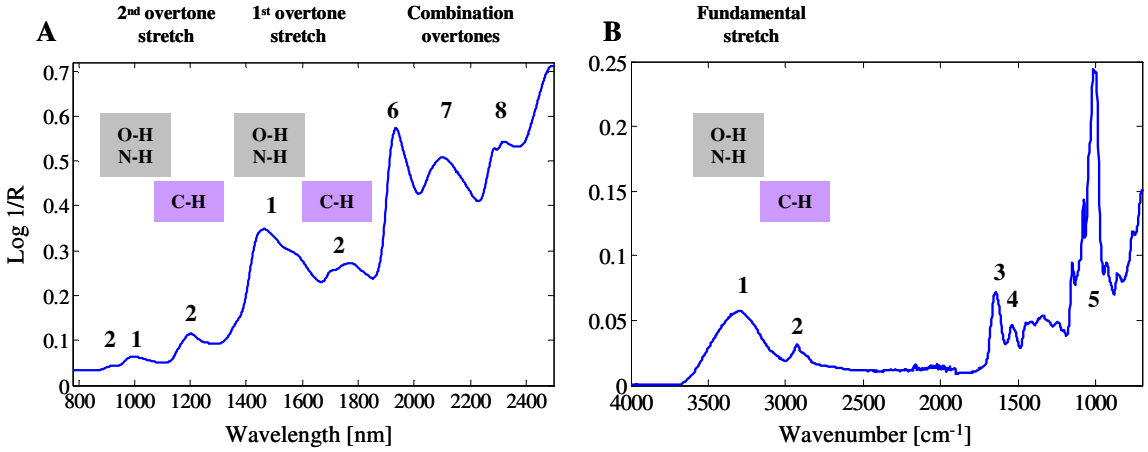


Figure 2.4: **A**: NIR spectrum of a typical bread wheat flour. **B**: The corresponding IR spectrum of the same wheat flour sample. Selected vibrational bands assigned: **1**: O-H and N-H stretch, **2**: C-H stretch, **3**: Amid I at 1640-1660 cm^{-1} (C-N + C=O stretch), **4**: Amid II at 1530-1540 cm^{-1} (N-H bend + C-N stretch), **5**: C-O and C-N stretch, **6**: O-H combinations and N-H combinations, **7**: Amid combinations, **8**: C-H combinations.

The intensity and exact position of the absorption bands in IR and NIR are positively correlated to the change in dipole moments by absorption of the radiation. The dipole moment in turn is influenced by H-bonding, hence changes in pH and temperature is readily observable in NIR. The very broad and overlapping peaks in NIR spectra thus originates from both truly overlapping overtones and combination tones as well as distributions of different local molecular environments (more or less H-bonding) giving rise to broad peaks even for rather simple molecules such as water.

In Fig. 2.5 the reason for the success of near-infrared transmission (NIT) spectroscopy is unfolded. NIT spectra of the major constituents of wheat; starch, gluten (protein), fat and water are superimposed with typical spectra of wheat grain and wheat flour. The NIT region form 850 to 1050 nm covers the 3rd overtone stretch of C-H bonds and 2nd overtone stretch of O-H and N-H and due to the relative low absorbance (see Fig 2.4A) the sample thickness can be chosen quite freely up to a few centimetres. Large amounts of water in the samples may however influence the result.

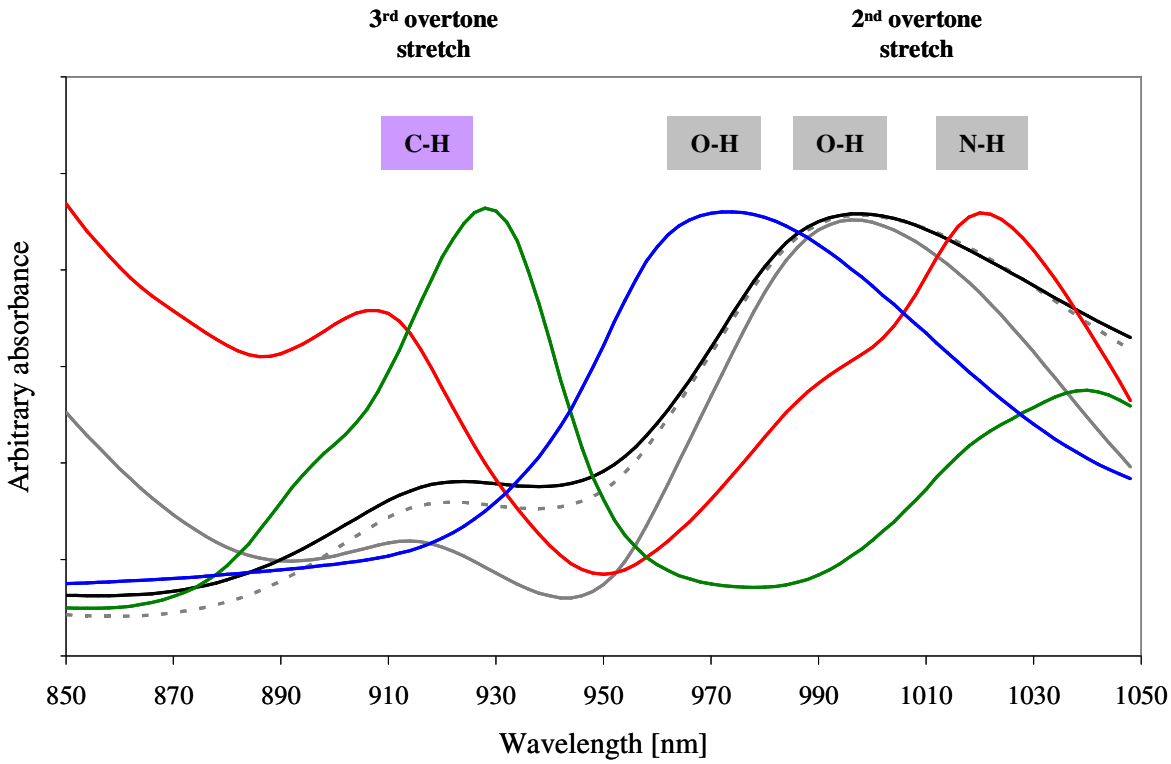


Figure 2.5: NIT spectra wheat grain (—), wheat flour (---), wheat starch (—), wheat gluten (—), plant oil (—), water (—). Modified from Pedersen (2002).

One of the great advantages of near-infrared spectroscopy (NIRS) in relation to food and feed analysis is the overwhelming amount of information present. Not only does it present quantitative information of O-H, N-H, C-H, S-H and C=O bonds in the irradiated samples, NIRS also provides an excellent tool for observing changes in the chemical conformation. This is because subtle changes in the chemical environment around those bonds influence the anharmonicity of the fundamental vibrations as well as the potential dipole moment change when a photon is absorbed. This may be looked upon a disadvantage, since these phenomena may devaluate Beer's law that states a linear relationship between absorbed light at concentration. However on the contrary the wealth of information can be readily extracted when used in relation with chemometric technology (Munck, 2005).

Multivariate NIR calibrations have been shown to work extraordinary well in agricultural product for fast determination of major as well as minor constituents (Miralbés, 2004, Williams and Norris, 2001). Protein, gluten, starch and moisture content of cereals are nowadays only measured by NIT based

on global calibrations, while minor constituents and properties such as individual amino acids determined independently of crude protein (Fontaine et al., 2002, Rubenthaler and Bruinsma, 1978) for feed production, rheological properties (Miralbés, 2004, Dowell et al., 2006b), gluten composition in terms of glutenin and gliadin contents (Wesley et al., 2001), starch damage (Osborne and Douglas, 1981, Osborne et al., 1982) and bread quality in terms of loaf volume and crumb structure (Delwiche and Weaver, 1994) are now being explored.

The low-cost, speed, accuracy and non-destructiveness as well as stability, operator safety (no solvents) and ease of use (flat learning curve) of NIR technology makes it a versatile technology for quality screening of al sorts of agricultural products with respect to optimal end use, value and price settlement as well as in breeding (Osborne, 2006).

2.5.2. Nuclear magnetic resonance (NMR) relaxometry and diffusimetry

Nuclear magnetic resonance is a versatile technique which can be used in various ways. The phenomenon utilised was discovered by Felix Block and Edward Purcell in 1946 for which they won the Nobel Prize in physics in 1952 (Brown, 1995). Depending on the instrument, i.e. magnet size, type of probe and geometry an endless variety of chemical and physical phenomena can be observed and analysed, e.g. diffusion of water in a complex matrix, brain scanning for tumours, chemical structure determination and metabonomic studies of complex biological fluids. The studied phenomena is determined only by the experimental setup and radio frequency pulse sequence and acquisition used (Callaghan, 1991). In this thesis however, only relaxometry and diffusimetry will be covered briefly as it was used to study water and fat compartmentalisation in dough and bread in Paper III and IV.

^{1}H, ^{13}C, ^{17}O and ^{31}P are nuclei which act like small magnets with spins aligned parallel or anti parallel in a magnetic field and may be probed by NMR. The stronger the magnetic field the more magnetic is the sample. In this study only ^{1}H-NMR has been used.

By applying a radio frequency equal to the so called Larmor frequency[6] the spins are perturbed (exited) and the transverse bulk magnetisation can be measured by a coil while the spins relax back to equilibrium, i.e. low energy state. The frequency and relaxation speed depends on the configurations of the protons and may be utilised for characterisation of a given sample. The frequency pattern of the sample may be studied by recording a free induction decay (FID) followed by Fourier transformation to study the spectrum in which the chemical shifts of protons can be assessed. This is known from structure determination of pure chemicals and "omics"-studies of complex materials. Recording spectra however, require a strong magnetic field in order to obtain an appropriate resolution as it is based on very small differences in the Larmor frequencies for nuclei relative to local shielding by the surrounding electrons (Brown, 1995).

Relaxation experiments in which spectral information is not of particular interest may be studied in fields as week as the earths magnetic field (0.05mT), using simple systems such as the Terranova-MRI (Magritek, Wellington, New Zealand). However these experiments may equally well be determined in high-field instruments (>7 T) employing advanced superconducting electromagnetically induced fields (Paper IV).

Relaxation takes place via two relaxation processes, spin-lattice and spin-spin relaxation. The lattice is the local surroundings of the exited proton. Spin-lattice relaxation takes place by energy exchange between the excited nuclei and the local molecular lattice field with energy transition levels equal to the Larmor frequency. Thus relaxation causes a very small heating of the sample. Spin-lattice relaxation is also called longitudinal relaxation. Spin-spin relaxation is caused by exchange of energy between exited protons with equal Larmor frequencies and thus disperses the energy. Spin-spin relaxation is also called transverse relaxation. Various relaxation experiments can be made to observe the different relaxation phenomena. Most common are the inversion recovery (Vold et al., 1968) and the Carr-Purcell-Meiboom-Gill (CPMG) experiments (Carr and Purcell, 1954, Meiboom and Gill, 1958) in

[6] The Larmor frequency (ν) is proportional to the magnetic field strength (B): $\nu = \gamma B/2\pi$, where, γ, is the gyromagnetic ratio which for ^{1}H is $2.675 \cdot 10^8$ $s^{-1}T^{-1}$. Protons precess at the Larmor frequency and in order for absorption of the RF pulse to take place *resonance* must be present, hence the name: "nuclear magnetic resonance".

which the longitudinal and the transverse relaxation can be explored and the characteristic relaxation time constants T_1 and T_2 respectively can be calculated for the components in the matrix studied. Depending on the actual instrumental setup NMR can be utilised in countless ways to study these spin states both spatially and dynamically in order to understand processes in complex materials and living tissues (Callaghan, 1991).

In Paper III a series of CPMG relaxation experiments were conducted during rising, baking and cooling of a small dough sample within a variable temperature probe of the 23.2 MHz (0.545 T) Maran benchtop pulsed NMR Analyser (Resonance Instruments, Witney, UK). The experiment was similar to the NMR-baking experiments of Engelsen et al. (2001), however this experiment included cooling of the sample. The raw data of one sample is shown in Fig. 2.6. The relaxation baking profile were used as is for predictions of bread quality in Paper III, as it has previously been shown that the entire decay curves works well in exploratory and prediction studies of food matrixes using PCA and PLS (Engelsen et al., 2001, Micklander et al., 2002, 2003).

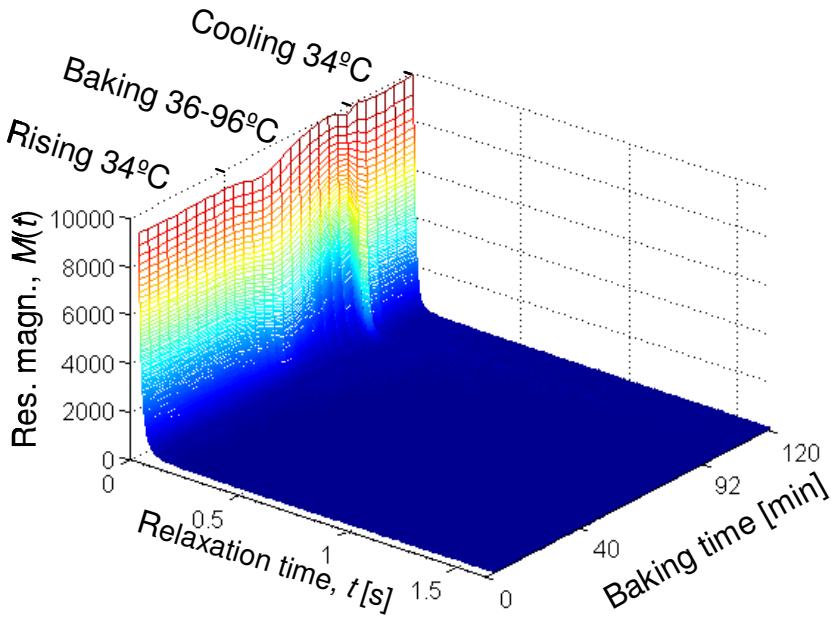


Figure 2.6: T_2 relaxation-baking profile of a dough sample (Paper III).

The possibilities to interpret and quantify the water and fat T_2-components by fitting the exponential components are covered in Chap. 4. More advanced experiments involving diffusion-relaxation correlation studies were

also made in dough samples by combining the pulsed gradient stimulated echo (PGSTE) sequence with CPMG acquisition (Paper IV). The labile nature of conventional 2D Laplace transformation of such 2-dimensional exponential decays was addressed and a new chemometric method for analysing such data was proposed in Paper IV, the PARAFAC-Laplace decomposition.

2.6. Exploratory data analysis

As representative sampling is physically democratic (Chap. 2.1) so is multivariate exploratory data analysis with respect to the data. Every variable and every sample count and those who collectively covary or agree the most has the most to say. They determine the directions in the data – the common references or underlying phenomena. Covarying as well as outlying samples and variables are easily recognised in scores and loadings plots. Samples in the same group agree on the values of the trend-setting variables and are easily recognised in a scores plot. Samples or variables which stand out from the crowd are either very interesting or just simple outliers, as a result of erroneous sampling, measurements or laboratory failures (e.g. wrong labelling). Exploratory data analytical tools such as principal component analysis (PCA) (Hotelling, 1933, Pearson, 1901, Wold, 1966, Wold et al., 1987) and partial least squares regression (PLS) (Geladi and Kowalski, 1986, Kowalski et al., 1982, Wold, 1982) for bi-linear data matrixes and parallel factor analysis (PARAFAC) (Bro, 1997, Harshman, 1970, Carroll and Chang, 1970) for multi-linear data are central to chemometrics. These tools provide virtually assumption-free analyses of the major trends and outlier detection in the data. The results are readily visualised in various plots for easy interpretation by the human brains cognitive apparatus (Munck, 2005). Although the soft multivariate models (Wold, 1975, 1982) mentioned above have certain advantages, traditional univariate models for analysis of variance (ANOVA) still has a role to play and may provide complementary as well as confirmatory information (Paper II).

The mathematical and statistical tools in science are an integral part of the solutions and the knowledge derived. Without comprehensive knowledge of the methods 'out there' and their use, the risk of drawing wrong conclusions is imminent. However, it is often the local scientific tradition (Kuhn, 1970) which governs the choice of data analytical methods used, making the new ideas of soft multivariate modelling in chemistry a struggle for those involved in the early days (Esbensen and Geladi, 1990, Geladi and Esbensen,

1990). The food science community and the food industry has always been working with very complex system, hence adapting new methodologies to solve current problems has always been an integral part of the work, e.g. the Farinograph, texture analysis and fast determination of protein and moisture by chemometric technology are such examples.

In general modelling data is all about separating structure from noise (Eq. 2.5):

Observed data = structure + noise (Eq. 2.5)

The structure part is the model. The reason for modelling data is to bring insight into the system under observation and to answer various questions we have about a system: What are the trends and relations in the data? Does the model chosen explain the empirical data reasonably well? Is our *a priori* hypothesis valid? The noise part is the leftovers and is often used to validate the model and is as such part of the model. Does the noise part actually look like white noise, when plotted or is there still structure left in the noise? Or is extreme behaviour present in the sample or variables measured indicating outliers or malfunctioning equipment?

Eq. 2.5 actually holds for all for all models of empirical data. The models used in this work (Paper I-IV) all conform to this basic contemplation. In the following PCA, PLS and ANOVA models and their use with respect to Papers I-III will be summarised. The PARAFAC model is presented in Paper IV. In addition different spectral data pre-treatment procedures will be covered. These are also part of any model, but are usually treated separately as the pre-treatment as the name implies is the first part of a stepwise operation followed by additional modelling, e.g. by PCA or PLS. In Chap. 4 multi-exponential modelling of NMR relaxation and 2D-diffision-relaxation data will be covered in relation to 2D PARAFAC-Laplace decomposition (Paper IV).

2.6.1. Principal component analysis (PCA)

Principal component analysis (Hotelling, 1933, Pearson, 1901, Wold, 1966, Wold et al., 1987) decomposes a centred data matrix **X** into a structured part and a random part, i.e. noise (Eq. 2.5). Strictly **X** need not be centred for the analysis, however, the purpose of the analysis is to find the differences and

similarities between objects and simultaneously determine which variables are important for the decomposition. Hence, centring and eventually scaling of the variables across all samples prior to PCA is usually implied.

In PCA, **X** is projected onto a new set of *N* latent variables chosen so that each latent variable, $\mathbf{t}_n$ is orthogonal, i.e. uncorrelated, to each other and successively describe as much of the variation as possible. A new coordinate system spanned in the **X** space is defined by *N* unit vectors $\mathbf{p}_n$'s formed by the analysis. Thus PCA is a reduction of dimensions from the number of variables, *J*, in **X** to *N* underlying virtual variables describing the structured part of data. The PCA model for **X** is:

$$\mathbf{X} = \mathbf{TP}' + \mathbf{E} = \mathbf{t}_1\mathbf{p}'_1 + \mathbf{t}_2\mathbf{p}'_2 + \ldots + \mathbf{t}_N\mathbf{p}_N + \mathbf{E} \qquad \text{(Eq. 2.6)}$$

$\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_N]$ are the scores and $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N]$ are the loadings. *N* is the number of components, underlying structures or effective rank of the matrix. **E** is the residual matrix, the part of data, which was not explained by the model, **TP'**. **E** is also called noise and error matrix and has the same dimensions as **X**. **E** is often used as a diagnostic tool for identification of outlying samples and/or variables. In Appendix B the non-linear iterative partial least squares (NIPALS) algorithm for calculation of the principal components is presented with a numbers on example of a PCA analysis with plots of scores and loadings for interpretation.

2.6.2. Partial least squares projections to latent structures (PLS)

The first comprehensive tutorial was published by Geladi and Kowalski in 1986. This paper offers what has become the classical introduction to PLS by a consecutive and graphical presentation of multiple linear regression (MLR), PCA, principal component regression (PCR) and PLS. By August 2007 this paper has been cited 1441 times in source items indexed within Web of Science.

PLS was first used on chemical data by Kowalski et al. (1982) and has become a standard tool in chemometrics by turning collinear data (e.g. spectral data) into an advantage rather than a problem (Wold et al., 1984).

PLS serves two major purposes. One is to establish a regression model between coherent **X** and **Y** data, e.g. NIT and flour quality parameters meas-

ured, in order to predict or estimate future **Y** (response) quality data by only measuring **X** (predictor) and applying the already establish regression model:

$$\mathbf{Y} = \mathbf{XB} \qquad \text{(Eq. 2.7)}$$

B is the regression coefficients estimated during the calibration of the PLS. The other purpose is to identify information in **X** relevant for **Y** in an exploratory quest to understand the relation between **X** and **Y**, if any exists. While PCA was an unsupervised analysis only guided by the variance of **X** itself, PLS iteratively models both **X** and **Y** while exchanging scores in order to explain as much variation as possible of both blocks simultaneously and maximise correlation between **X** and **Y** scores. Thus the decomposition is supervised by the variation in **Y**. If only one response variable exists, the model becomes slightly simpler:

$$\mathbf{y} = \mathbf{Xb} \qquad \text{(Eq. 2.8)}$$

The PLS models are calculated using the NIPALS algorithms, PLS2 and PLS1 respectively, similar to the one used in PCA (Appendix B) (Bro, 1996, Esbensen, 2000). PLS1 with only one response variable, **y**, is very fast as no iterations are necessary for optimisation. Thus PLS consists of two 'PCA-like' models representing the outer relations:

$$\mathbf{X} = \mathbf{TP}' + \mathbf{E} \qquad \text{(Eq. 2.9)}$$
$$\mathbf{Y} = \mathbf{UQ}' + \mathbf{F} \qquad \text{(Eq. 2.10)}$$

X is decomposed into **T** and **P** like in PCA, however under the influence of **Y**, hence **T** and **P** are called PLS-scores and PLS-loadings. **E** is the residual matrix of **X**. Similarly **Y** is decomposed into **U** = **Y**-scores and **Q** = **Y**-loadings. A set of loading weights, **W**, are also calculated when **X** is projected onto **U** for each component. The regression coefficients are calculated from **X**- and **Y**-loadings and loading weights:

$$\mathbf{B} = \mathbf{W}(\mathbf{P}'\mathbf{W})^{-1}\mathbf{Q}' \qquad \text{(Eq. 2.11)}$$

For interpretation purpose **P**, **W**, **Q** and **B** are worth inspecting. High values generally mean high importance. **W** is the covariance of each variable in **X** to the variation in **Y**. **Q** are the **Y**-variables relevant to **W**. **P** contains both **Y**-relevant as well as **Y**-irrelevant information, which may be difficult to interpret. **B** is the direct link between **X** and **Y**, thus making these very important. However, interpretation may be an erratic affair since PLS is modelling both variations of interest and interferences in **X**.

2.6.3. Variable selection

While full scale modelling may be erratic, variable selection may reduce the number of variables for better modelling and interpretation. Several methods exist in which variables are selected collectively or individually in all possible combinations in order to automatically select the optimal set of variables for prediction purposes and interpretation purposes. Genetic algorithms (GA) (Leardi, 2001), Jack-knife cross validation (Martens and Martens, 2000) and interval PLS (iPLS) (Nørgaard et al., 2000) are examples of such automatic iterative methods which help researchers to focus on the important parts of multivariate data and simultaneously filter out regions or variables which contain only noise or interfering information. However due to the sheer number of variables in modern data acquisition, the risk of overfitting should not be underestimated. In Paper I, iPLS was used as an algorithm to select the most appropriate way to present single kernels in the near-infrared light beam. Although the input data was a number of concatenated and thus already separate NIT spectra blocks originating from different sample presentations, the algorithm proved well suited as a data block selection tool.

2.6.4. Multiblock PLS (MBPLS)

In Paper III an opposing method to variable selection was used to evaluate whether conceptually meaningful blocks of coherent data would contain complementary data for better prediction of wheat and bread quality. The multiblock PLS (MBPLS) method did not differ from the PLS model presented above apart from **X** being a set of **X** blocks ($\mathbf{X}_1$ to $\mathbf{X}_7$) containing different kinds of information, e.g. NIT on grain, NIT on flour, NIR on flour, IR on flour etc.

The fundamental model is actually the same and the prediction performance is exactly the same as if all conceptually separate blocks were merged in one

X block. The advantage is merely the ordering of data and the possibilities to interpret the outcome in a comprehensive way. Having a common NIPALS origin various multiblock PCA and PLS models and algorithms have been developed, optimised and changed over the years (Westerhuis et al., 1998) and will not be covered here. In this work the multiblock PLS as described in Qin et al. (2001), Westerhuis and Coenegracht (1997) and Westerhuis et al. (1998) with deflation of **X** using super scores has been used. The basic algorithms for calculation of the models are available from www.model.life.ku.dk.

Blocking can improve the interpretability when a large number of variables can be divided into conceptually meaningful blocks (Westerhuis et al., 1998). The blocks may contain information from different instrumental techniques, such as NIR and IR spectra on the same material as in Paper III and in Brás et al. (2005) or from the same instrument but at different stages in the process as in Paper III, where NIT was used on grain lots and flour lots, or at specific time points in a process (Qin et al., 2001, Choi and Lee, 2005). The blocking of data is clearly problem dependent and based on the knowledge of the problem at hand. Multiblock modelling can be used in various ways either exploratory in order to identify important data sources or processing steps (Paper III, Westerhuis and Coenegracht, 1997, Brás et al. 2005) or in process monitoring for process control and fault detection (Qin et al., 2001, Choi and Lee, 2005). The MBPLS model used in this work may be calculated directly from ordinary PLS (Qin et al., 2001, Westerhuis et al., 1998). One of the problems with this current MBPLS approach is that the deflation of the blocks is based on the super-scores, thus effectively mixing up information in the blocks in the deflation step. Westerhuis and Smilde (2001) suggested deflating only **Y** by modifications of the algorithm to make the block weights, scores and loadings more interpretable.

2.6.5. Weighting and scaling variables and blocks in MBPLS

Weighting and scaling of individual variables and blocks requires some attention as erroneous scaling/weighting may lead to blocks and/or individual variables having to much or too little influence on the model calculated. Within each block data are either centred, standardised (unit variance of each column) or range scaled or subject to a combination of those. Centring removes the offset and is often a standard procedure in multivariate data analysis, as centring focus the analysis on differences and similarities of in-

dividual samples rather than the overall level. Standardising, i.e. dividing each entry of each variable with the standard deviation of the entries, evens out the difference in scale or unit of each variable. This is often used where variables are of different origin and thus of different scales, e.g. protein concentration, water absorption, falling number determined in a flour sample. Scaling in specific ranges may also be used, e.g. 0 to 1. In spectra or spectra like data, such as near-infrared spectra, NMR relaxation decay curves and chromatograms, were low intensity usually means low importance, scaling or standardisation may have detrimental effect by inflating noisy, low intensity variables.

After weighting the variables within each block, each block should be weighted in order to assure appropriate influence of each block. The number of variables in each block may vary immensely from a hand full to thousands. Thus in order to induce equal importance of those blocks they are usually normalised to have block unit variance, i.e. the sum of squares of the entire block is adjusted to one, by multiplying each entry with the appropriate constant. The MBPLS had comprehensive handles for this operation, while this could have been done equally well on ordinary multivariate calibration software.

2.6.6. Validation

Validation of PLS models can be done in several ways. The goal is to establish the appropriate model complexity which is usually done by evaluating the **Y** prediction error. The point at which the prediction error is not decreasing anymore by adding another component is the model complexity. The prediction error is evaluated by the root mean squared error (*RMSE*) of predicted versus measured y-values. This may be done either based on the calibration (*RMSEC*), cross validation (*RMSECV*) (Paper I and III) or on an independent test set (*RMSEP*). The latter is of particular importance when the prediction error is estimated with regards to new samples to be predicted based on the model (Esbensen, 2000).

2.6.7. Analysis of variance (ANOVA)

Analysis of variance (ANOVA) developed by R.A. Fischer in the 1920's is used in connection with a specific experimental design where controlled factors are varied and their effect observed on the response variable measured (Fischer, 1950 cf. Youden, 1951, Skovgaard, 1996). The model is chosen

in accordance with the factors varied and the actual design of the experiment. This is by no means simple and involves a number of assumptions regarding the samples, the model parameters and the distribution of the noise. In addition the actual model and the tests made depend on a number of things, including missing values, random effects, blocks etc. On the other hand ANOVA belongs to the inferential statistics. Given the correct model and valid assumption, conclusion can be made regarding the factors varied associated with a probability of being wrong by coincidence. Hence the focus of the ANOVA is the factors varied – not the similarities of individual samples as in PCA or prediction of new samples as in PLS. In Paper II a 3-factor ANOVA model was used to determine the effect of growth year, location/cultivar and catch crop on a number of variables measured in a bread baking process. On top of this, the effect of sorting the bulk crop was investigated on the same material from the same growth years and locations. The model used conforming to Eq. 2.5 separating structure from noise:

$$X_{ijk} = \mu + A_i + B_j + \gamma_k + \varepsilon_{ijk},\ i = 1,2;\ j = 1,2;\ k = 1,2,3 \qquad \text{(Eq. 2.12)}$$

X_{ijk} is the measured response, e.g. dry matter protein, μ is the overall mean value, A_i is the random effect of growth year 2003 or 2004, B_j is the random effect of location Arslev or Kiel, γ_k is the fixed effect of catch crop, winter vetch, fodder radish or no catch crop and ε_{ijk}, is the residual unexplained residual. In the test of fractionation, γ_k is the fixed effect of fraction 1, 2 or 3. For the model parameters apply: $\sum\gamma_k = 0$ and A_i, B_j and ε_{ijk} are independent and $N(0,\sigma_A^2)$, $N(0,\sigma_B^2)$ and $N(0,\sigma^2)$ respectively. All the lots involved in this investigation are independent. Eq. 2.12 omits the possible random interaction effects of A_iB_j, $B_j\gamma_k$ and $A_i\gamma_k$ as these are of no interest in the specific investigation[7].

[7] Including the interaction terms in Eq. 2.12 would complicate the investigation immensely as year and location are random factors. The actual F-tests to be made in the ANOVA regarding the main random as well as fixed effects would depend on which of the random interaction effects were considered significant (Skovgaard, 1996). The conclusions made from such a complicated and elaborate task would be doubtful anyway due to the low number of experiments, i.e. low degrees of freedom, and the number of variables tested. Thus choosing the model with only main effects implies that the variance of the interaction terms are all zero, hence insignificant, $\sigma_{AB} = \sigma_{A\gamma} = \sigma_{B\gamma}, = 0$.

Given all the above is approximately valid, the hypotheses that $\sigma_A^2 = 0$, $\sigma_B^2 = 0$ and $\gamma_1 = \gamma_2 = \gamma_3 = 0$ can be tested. Probabilities (P-values) for there being no effect of the factors are calculated according to the ANOVA. Hence a small P-value indicates that a given hypothesis regarding the absence of effect is unlikely to be true. By convention P-values smaller than 0.05 are called significant and marked with a '*', while stronger degrees of significance below 0.01 and 0.001 and marked with '**' and '***' respectively. Only the P-values are reported in Paper II.

2.6.8. Pre-processing spectroscopic data

The concept of pre-processing is to remove irrelevant systematic variation from the data without destroying the chemical information too much prior to multivariate modelling. The goal of pre-processing is either to improve the subsequent calibration model to gain better robustness and prediction performance (e.g. reduce the number of PLS components and/or the prediction error) or to improve the interpretability of the models (e.g. improve scores, loadings and regression coefficients which may appear erratic in complex models). In altering the data by any rank reducing pre-processing method, there is a risk of over-doing it – and actually destroy the data. Pre-processing is thus not the answer to all problems in data analysis. Often it is an iterative trial and error process in order to find a suitable method and its associated parameters to a current problem. The performances of the methods tested are evaluated towards an objective goal such as the prediction error, the number of model components and/or the fingerprint (plot) of the residual. However, the number of different pre-processing methods and alterations published is quite overwhelming (Gabrielsson and Trygg, 2006) and not one method can be said to be objectively superior to the other, but depends on samples, spectroscopic methods and available prior knowledge. The slow implementations in commercial software packages as well as patent rights make new methods less available for general applications.

In light spectroscopy there are some acknowledged physical challenges, that can be met by hard modelling prior knowledge. Using spectroscopy to determine the concentration of a chemical constituent, Beers law usually applies (Chap. 2.5.1) – stating a linear relation between the absorption of light and the concentration. The path length however, is not necessarily constant and contains both and additive component and a multiplicative component. This is true for almost all types of samples apart from transparent one phase

solutions. When multiple phases are present such as particles in air, slurries and emulsion the light passes through different materials and is refracted many times on its travel and thus does not experience the specified path length. It follows a completely different route depended on the number and character of interfaces encountered on the way. The number of phase transitions and thus refractions determines the number of deflections from the set path. Samples containing small particles are thus thought to experience longer path length than samples of the same material containing large particles. This may lead to an additive effect in which spectra of the same material is displaced constantly over the entire spectrum. In addition – photons travelling longer in the material are more likely to be absorbed, thus a multiplicative effect is observed.

Consider photons at two energy levels, E1 and E2. E1 has an energy that is not very likely to be absorbed, while E2 is very likely to be absorbed – given they travel the same distance. That is due to the chemical composition of the material in which transition level E2 is present more than E1. Changing the path length a little by selecting differently sized or differently packed particles will lead to a proportional change in absorption at both wavelengths. In the spectra, this is observed as a bigger change in absorption at E2, than at E1. That is the multiplicative effect.

In multiphase systems diffuse scatter will be observed caused by reflection of the light on the interfaces between phases. The smaller the particles – the more scatter is observed. This effect will result in apparent extra absorption in transmission mode as the reflection never arrives at the detector, while in reflection mode the prerequisite for measuring anything is the diffuse reflection from the sample. Thus in reflection mode bigger particles will lead to actual more absorption and vice versa for small particles (Dahm and Dahm, 2001).

In order to circumvent these physical effects dedicated methods have been developed. Multiplicative signal correction (MSC) sometimes called multiplicative scatter correction due to its origin in near-infrared scatter correction (Geladi et al., 1985) has become a reference to which new methods are compared (Martens et al., 2003, Pedersen et al., 2002a). Martens and Stark (1991) introduced an extended version incorporating potential prior knowledge such as pure component spectra - later extended even further with wave-

length dependencies and squared spectra. The extended multiplicative signal correction (EMSC) (Martens et al., 2003) is a pragmatic solution in which extensions may be added at will. Standard normal variate (SNV) with or without subsequent de-trending provide similar results as the MSC method (Barnes et al., 1989). The above mentioned methods have the advantage of preserving the shape of the spectra, which may be an advantage for interpretation. Spectral derivation on the other hand does not preserve spectral shape, but effectively remove both additive (1st derivative) and multiplicative effects (2nd derivative) while preserving relevant information (Savitsky and Golay, 1964). Delwiche (1995) combined the Savitsky-Golay derivation and the MSC, in that order, providing superior robustness and prediction error for prediction of protein content of single seeds using single-kernel near-infrared spectroscopy. This information was utilised and confirmed by Nielsen et al. (2003), Pedersen et al. (2002a) and in Paper I for prediction of protein in single wheat kernels.

In Fig. 2.7 the effect of four different pre-processing methods is demonstrated graphically with respect to the shape of the spectra and the information discriminating the wheat lots involved. Fig. 2.7A shows the raw near-infrared transmission spectra of thirty-two wheat grain lots investigated in Paper II and III. The spectra are coloured according to growth year and location as well as winter cultivars. Clearly the spectra do not differ very much in shape but rather in general level due to different physical appearances of the grains. However, the physical appearance actually contributes to the discrimination of the lots. In Paper II the quality fractionated lots also resulted in systematically different raw NIT spectra, thus indicating that internal quality can actually be seen on the outside of the grains. In Fig 2.7B the spectra are centred prior to principal component analysis in Fig. 2.7C. The growth years, 2003 and 2004, are clearly reflected on the first PC, while the second PC reflects the location, Aarslev vs. Kiel – disregarding the groups 04Se and 04Wi, which by design are deviating from the other samples. In Fig. 2.7D the spectra was corrected with MSC. This changed the spectra quite dramatically and at the same time transformed one of the winter cultivars (pent) into an outlier on the first components in the corresponding PCA plot (Fig. 2.7F) and similarly the bread wheat lots (Bre1 and Bre2) from Sejet on the second PC. The growth years and locations have switched positions and are no longer as clear as in Fig. 2.7C. By using EMSC incorporating wavelength dependencies (Fig. 2.7G, H and I), the year and location groups

are better discriminated in the PCA plot and those outliers are not as pronounced as in Fig. 2.7F. Taking the 2nd derivative in Fig. 2.7J indeed changed the shapes of the spectra and the subsequent centring and PCA score plot. By comparing the centred data in the middle column it is evident, that discriminations between the samples are possible, but it also elucidate that the important spectra regions depend on the pre-treatments and thus affects the score plots of the first two principal components.

In Fig. 2.8 five different strategies for pre-processing the same data as in Fig. 2.7 were compared by subsequently fitting a PLS regression model to protein content. A combination of 2nd derivative and MSC clearly reduces the rank of the data, however the EMSC provides the best prediction error using 9 PCs without encountering any local minima on the way. In similar ways the choice of EMSC as a general pre-processing for NIT of grain and NIT, NIR and IR of flour was made in Paper III. Experiments with a number of pre-treatment strategies for individual quality properties however may provide even better solutions (Delwiche and Graybosch, 2003).

As can be seen from the results in figure 2.7 and 2.8 pre-processing is not really a data wash or convenient filter prior to the actual modelling as the results as well as their interpretation is very dependent upon the method(s) employed. Hence pre-processing is certainly an integral part of the data modelling process. A formal development in that direction is the orthogonal projections to latent structures (O-PLS and O2-PLS) algorithms which incorporates the orthogonal signal correction (OSC) (Wold et al., 1998) pre-processing into the PLS algorithm (Trygg, 2002, Trygg and Wold, 2002, 2003, Verron et al., 2004). By removing all information not correlated, i.e. orthogonal, to $\mathbf{Y}$, the OSC filter and the O-PLS/O2-PLS has been shown to improve the interpretability of the PLS results (Samp et al., 2003, Svensson et al., 2002).

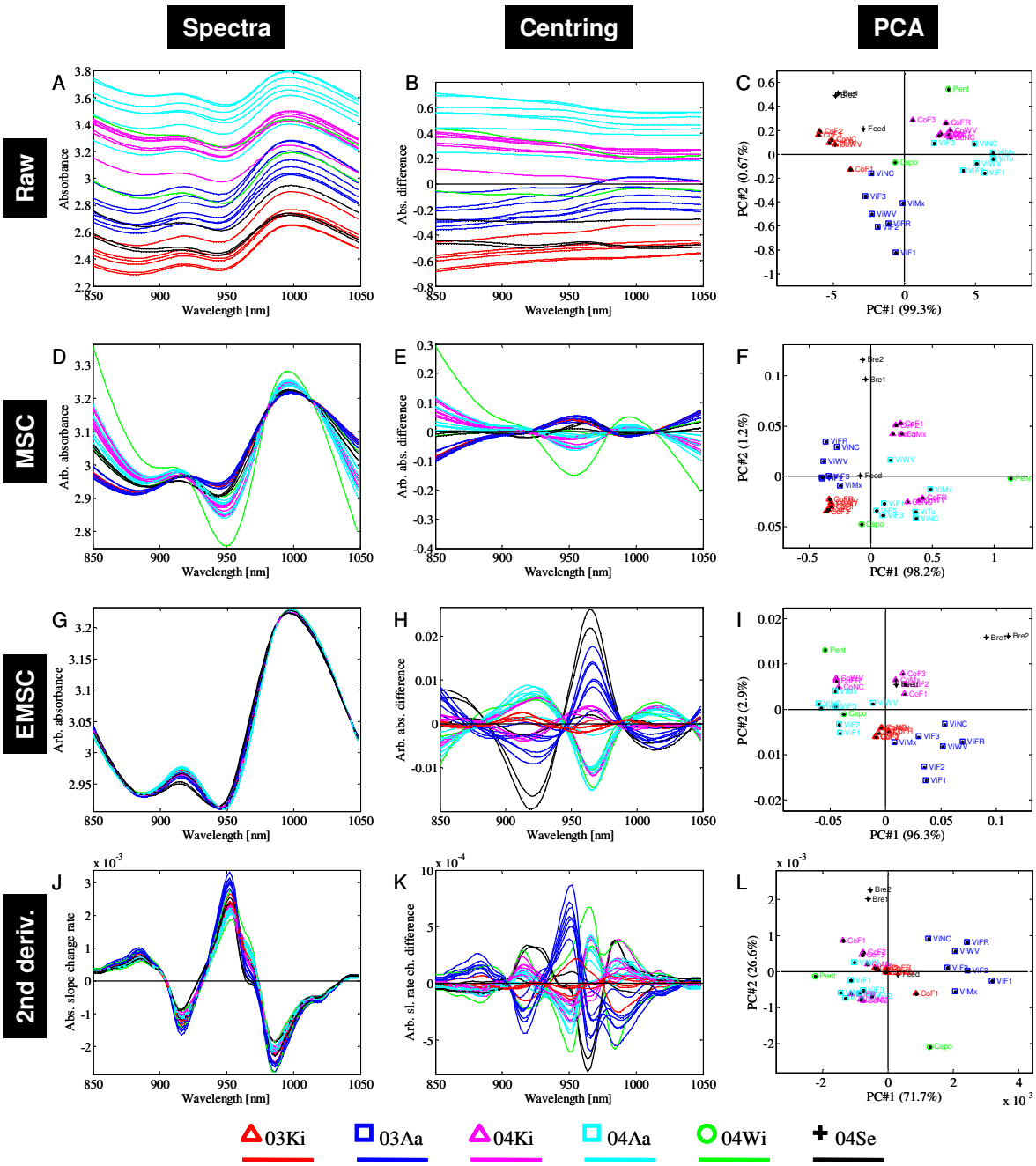


Figure 2.7: NIT spectra of thirty-two wheat grain lots pre-treated with either nothing, MSC, EMSC or Savitsky-Golay 2nd derivative (first column). The same spectra after subtraction of mean spectrum elucidate the major differences from average (second column). PCA scores of the centred data show major differences depending on the pre-treatment (third column).

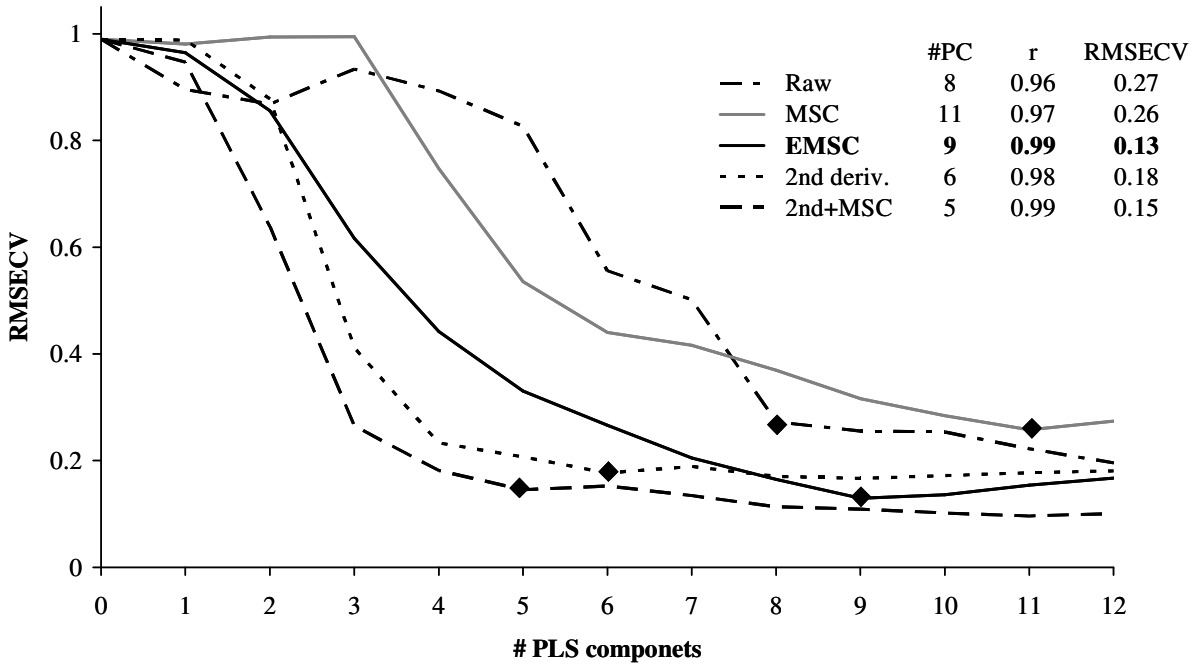


Figure 2.8: Root mean squared error of cross validation (*RMSECV*) of protein content in five PLS models using different pre-processing methods of near-infrared transmission spectra of wheat grains. Optimal model complexity marked by ♦ in the first local or global minimum in each model.

2.7. The PAT perspective revisited

The aim of presenting sampling, analytical methods and data analysis together in this chapter (above) was to emphasise that in all empirical sciences these individual fields are integrated. Data-analysis does not exist without measurements and measurements not without mass reduction and/or sample presentation. The conclusions drawn from scientific studies are inherently determined by the processes by which they were designed and obtained. The PAT initiative (USFDA, 2005) with regards to pharmaceutical processes is the blue stamp of the holistic scientific approach in all sciences and is foreseen to become the standard in future scientific and technological projects and developments.

3. Diversification and prediction of bread wheat quality

There is an increasing demand for flours meeting specific standards with regard to processing functionality as well as end product quality and uniformity (Mirablés, 2004). Thus improved abilities for fast determination and control of central quality parameters of both raw materials and during processing is very important. With reference to the inherent heterogeneity of bulk wheat (Paper I), processing by sorting to improve and diversify raw material according to quality poses a lucrative potential for providing dedicated uniform raw materials with narrow specifications (Papers II and III).

3.1. The holistic process analytical approach

In this work the holistic approach has been central to the endeavour of grasping the entire complex conversion of wheat from grain to flour to visco-elastic dough and finally to spongy bread crumb. The experimental process line is shown in Fig. 3.1 and consists of: 1. Raw materials diversified by agronomic measures and prior knowledge. 2. Post harvest diversification of selected wheat lots by mixing and subsequent single-kernel sorting according to a complex quality trait. 3. Milling. 4. Baking. The aim of the approach was first to evaluate the effect of sorting with regards to the quality of the resulting flours as well as the quality of the final bread products (Paper II). Secondly the possibility to predict the end quality based on multivariate spectroscopic measurements on bulk grains, flours and dough was investigated (Paper III). An inventory of methods for characterisation at all levels from grains to consumption was used in an exploratory attempt to embrace the variations of the materials as well as the process using process analytical technology (Fig. 3.1). “A process is generally considered well understood when (1) all critical sources of variability are identified and explained; (2) variability is managed by the process; and, (3) product quality

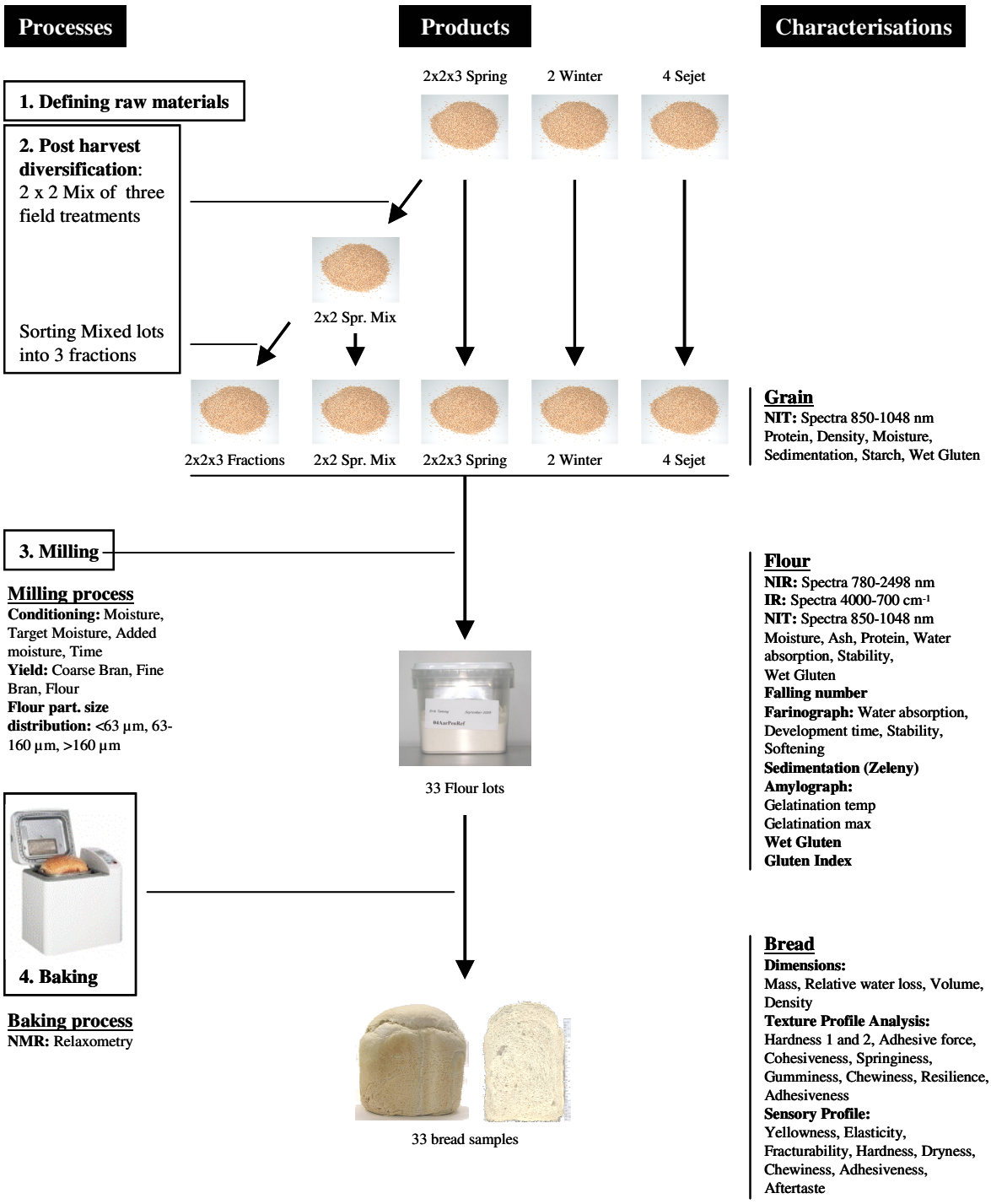


Figure 3.1: The process analytical approach to the baking process starting with, 1. Raw materials, 2. Improvements of raw materials by mixing and sorting followed by, 3. Milling the grain lots to flour and 4. baking breads (left column). The product developments can be followed in the middle column, while the data derived from the process by standard methods as well as by spectroscopic techniques are indicated in the right column. Some data closely related to the processes are shown in the left column along with the process steps.

attributes can be accurately and reliably predicted over the design space established for materials used, process parameters, manufacturing, environmental, and other conditions" (USFDA, 2004a).

3.1.1. A brief history on the experimental design

Single-kernel sorting was not originally intended to be a part of this work. The idea however developed from a rescue plan, in order to ensure appropriate diversity in an organic wheat material with an expected low diversity, to a central part of the wheat quality perspective. Confronted with the fall of the original idea comprising a full scale study involving field trials with several cultivars, fertilizing strategies and growth conditions this work suffered a minor set back. A new strategy had to be put in place without leaving the original process perspective. A search for material in storage from already performed field trials which fulfilled two major requirements was initiated. First requirement: The trial had to be of a certain size, so that a diverse material with respect to baking quality would be available. Second requirement: enough material from each experimental unit should be available so that the entire process from grain to brain could be performed. Apparently the local Interreg IIIA programme trials already running fulfilled those basic requirements, at least in terms of available material. However, the organic experiment focusing on the effects of various preceding catch crops for diversification of product yield and quality could be detrimental in terms of providing material of appropriate diverse nature. The organic fertiliser deficient environment might lead to low quality grain lots with low diversification, hence hampering the possibility to model and establish sufficient process understanding.

The solution came into place by utilising a new system under development by Bomill AB, Lund. Löfqvist and Nielsen (2003) had recently taken a patent on the TriQ sorting system for heterogeneous organic materials, such as grains. A laboratory scale single-kernel near-infrared (SKNIR) sorting device

had been developed for unsupervised[8] or supervised[9] sorting according to a complex quality trait. The system utilises the bulk heterogeneity (Paper I) as an integrated property of the population and as a source for bulk diversification.

By employing this technique to the organically grown material from the Interreg IIIA programme a new level of bulk diversification could be obtained. This diversification was utilised both for broader variation in the prediction study (Paper III) and for testing the effect of sorting as compared to the organic agronomic strategy of using catch crops to preserve and accumulate nutrients in the top soil for the subsequent wheat crops (Paper II).

3.1.2. Materials and methods

The raw material (Fig. 3.1 1st row) consisted of eighteen wheat lots of which fourteen originated form the Interreg IIIA organic field trial, while the remaining four conventional lots were generously donated by Sejet Planteforædling in Denmark. The lots from Sejet were included as references due to expected large diversity, i.e. one biscuit wheat (04SeBisc), one fodder wheat (04SeFeed) and two bread wheats (04SeBre1 and 04SeBre2). The fourteen organic wheat lots were from growth year 2003 and 2004 and grown in two locations, Kiel in Germany and Aarslev in Denmark. Two winter cultivars, Capo and Pentium were grown in Kiel and Aarslev respectively only in 2004 (04KiCapo and 04AaPent). Two spring cultivars, Combi and Vinjett were grown in Kiel and Aarslev respectively in both year 2003 and 2004 and subjected to three different preceding catch crops treatments; 1. No catch crop, 2. Winter Vetch (*Vicia villosa*), 3. Fodder Radish (*Raphanus sativus var. Oleiformis*) or Turnip (*Brassica rapa*). The last two used interchangeably both from the *Brassicaceae* family. Thus the spring cultivars resulted in twelve lots diversified by growth year and catch crop treatment (2 cultivars/locations × 2 growth years × 3 catch crops) (Fig. 3.1 1st row). This design was utilised in

[8] Unsupervised sorting refers to sorting according to predicted scores on latent variables from principal component analysis (PCA) or similar multivariate methods.

[9] Supervised sorting refers to sorting according to a predicted value of a specific quality parameter (e.g. protein content), which can be predicted from the multivariate signal based on a calibration method such as partial least squares regression (PLS-R) or similar methods.

Paper II for analysis of variance (ANOVA) with respect to the three factors (Chap. 2.6.7).

Half of the material from each of the twelve spring lots was mixed within each location and growth year resulting in four mixture lots (03KiCoMx, 04KiCoMx, 03AaViMx, 04AaViMx) (Fig 3.1 2nd row). The mixtures were sorted grain by grain into three equally sized fractions each (F1, F2 and F3) using the TriQ sorter (Fig. 3.1 3rd row). Hence, an alternative diversification within each year and location/cultivar were established and analysed with ANOVA in Paper II (2 cultivars/locations × 2 growth years × 3 TriQ fractions).

Only the organic lots were investigated in Paper II, while all available lots were included in Paper III. Two lots however were discarded. 03KiCoMx were discarded after fractionation, as too little material was left. 04SeBisc were discarded as this lot was unfortunately not submitted to conditioning with water prior to the milling. The missing conditioning step, mistakenly omitted due to high water content (16.3%), resulted in an inferior flour yield of only 58.3% and a markedly deviating particle size distribution as compared to the rest of the wheat lots.

3.2. Single-kernel diversity – a source for bulk diversification

By grabbing a few wheat grains it is quite clear, that each kernel is unique and differs in readily observable morphological characteristics such as size, shape and colour (Fig. 3.2). In Paper I a great variation in kernel mass and protein content was found among the investigated lots. The average dry matter protein contents were between 9.9% and 15.8% with kernel standard deviations of 1.1% to 2.5%. Delwiche (1995) showed similar variation and also that the protein content of single kernels was not correlated to their mass. Other morphological properties have been assessed by image analysis together with protein content, density and hardness on individual kernel of European wheats (Nielsen et al., 2003). In their study the single kernel weight varied from 24.1 to 69.3 mg and the protein content from 6.8% to 17 % dry matter, comparable to the findings in Paper I and Fig. 3.2. Nielsen et al. (2003) found that individual kernel density varied from 0.99 to 1.25 g cm^{-3}. Positive correlations between single-kernel protein content and kernel

density, protein content and apparent vitreousness[10], and density and apparent vitreousness were found with correlation coefficients around $r \approx 0.6$. Although single kernel hardness was positively correlated to apparent vitreousness ($r = 0.55$) no convincing correlations between hardness and protein, and hardness and density were found.

	1	2	3	4	5	6	7
W01 Pentium							
							
Mass [mg]	47.9	54.8	36.8	33.7	57.6	53.0	49.3
Protein [%dm]	9.7	10.9	10.6	9.3	10.2	11.3	9.4
W09 Pentium							
							
Mass [mg]	61.7	35.0	46.6	37.5	33.4	44.7	53.2
Protein [%dm]	12.9	10.9	11.7	10.1	9.7	11.1	11.3
S10 Vinjett							
	n.a.	n.a.					
Mass [mg]	47.3	48.3	24.2	46.4	39.3	43.1	29.3
Protein [%dm]	12.3	14.1	9.6	16.1	10.1	9.8	11.7

Figure 3.2: Heterogeneity in wheat lots with regards to appearance, mass and protein content visualised with seven random wheat kernels from three different wheat lots analysed in Paper I.

In the milling industry the quality of the raw material is often assessed by the above mentioned parameters and the value as well as payment is often settled based on protein content. A high test weight, i.e. the bulk measurement for density, is important for high flour yield while vitreousness and hardness is assessed for optimal milling processing determining flour particle size distribution and amount of damaged starch. Protein is thought to play a central role, as storage proteins surround and fill the spaces around the starch granules in the kernel endosperm entirely in hard wheats, while

[10] Vitreousness is an assessment of how glassy as opposed to how opaque a kernel appears. In Fig. 3.2 kernel no 4 in sample S10 Vinjett has a much darker colour and vitreous appearance indicating that the space between starch granules is completely filled by the protein matrix.

only partly for soft wheats (Huebner and Gaines, 1992). Greffeuille et al. (2006) discuss the complexity of grain hardness, vitreousness and protein content in relation to milling. In their study of near-isogenic wheat lines, hardness appeared to be a genetic factor, while vitreousness was primarily an environmental factor, although hard wheat generally appears more vitreous than soft wheat. Increased vitreousness had a positive effect on flour yield in soft wheat while a negative effect in hard wheat milling properties was observed. The protein content neither correlated to vitreousness nor hardness. These findings are supported by the low correlations found by Nielsen et al. (2003). While protein content is positively affected by growth conditions in more or less the same way as vitreousness their interdependency may in general be overemphasised, at least as a causal relation. Greffeuille et al. (2006) concluded that hardness may be related to adhesion forces between starch granules and the protein matrix, while vitreousness is related to endosperm microstructure or porosity.

Huebner and Gaines (1992) report that cultivars classified as either hard or soft wheats are themselves quite heterogeneous with respect to individual kernel hardness. The sizes of kernels were apparently larger when growing close to the middle of the kernel head as opposed to top and bottom. Their origin on the kernel head, hence the kernel size however was not correlated to hardness. Hardness was apparently correlated to a single gliadin fraction separated on RP-HPLC, hence related to the quality of the protein.

3.2.2 Single-kernel sorting

Post harvest utilisation of the bulk wheat heterogeneity or diversity appears to be lucrative. However, in order to sort the grains to form more uniform and higher value fractions a series of events should take place. A system suitable for sorting should have an ultra-fast positioning device, an ultra-fast determination of quality, an ability to keep track of the individual grains and a mechanism to place the grain into the appropriate receiving bin. An objective goal of such a system could be sorting at least one ton an hour, corresponding to $2.5 \cdot 10^7$ grains with an average mass of 40 mg each. Quality determination of one grain with subsequent delivery should take no more than 0.14 ms.

3.2.3. Sorting with respect to grain morphology

A starting point would be basing the quality determination on readily observable morphological features such as size, weight, density and colour. Nielsen et al. (2003) argue that positive but relatively low correlation between protein content, vitreousness and hardness to density is not enough for a quality sorting based on density grading. Yoon et al. (2002) also concludes that the increased flour yield based on size and hardness sorting was insufficient to outweigh the cost of sorting. In barley, Elfverson et al. (1999) found only small differences in starch and β–glucan content with respect to the size of grains. Large grains had a tendency of containing more β–glucan and less starch as compared to fractions containing smaller grains. Protein contents tended to be lowest in the intermediate size fractions. Differences between cultivars were much larger. Elfverson et al. (1999) cite others for observing increasing, decreasing and no effect on protein contents based on grain size sorting in wheat depending on the starting material. They also noted that wheat in comparison to barley show much larger differences in chemical composition between fractions of grains. Pasikatan and Dowell (2004) did improve protein contents marginally by sorting mixtures of high and low protein wheat using a dual wavelength filter (920 nm and 1660 nm) which was partly reported to sort according to colour and vitreousness.

3.2.4. Sorting with respect to internal quality trait

With dedicated spectrometers and setups it has become possible to record near-infrared spectra of single kernels and relating these signals to various properties by chemometric calibration models. Prediction of hardness (Delwiche, 1993, Maghirang and Dowell, 2003, Nielsen et al., 2003), vitreousness (Delwiche, 1993, Dowell, 2000, Nielsen et al., 2003) and protein content (Delwiche, 1995, 1998, Delwiche and Hruschka, 2000, Nielsen et al., 2003) has been assessed by single-kernel near-infrared (SKNIR) spectroscopy. While protein content appears to be well determined by SKNIR, hardness and vitreousness calibrations are more difficult, probably due to large variations in the reference methods. Vitreousness is basically a subjective assessment (Dowell, 2000, Nielsen et al., 2003) and hardness determination is very dependent on the method e.g. accuracy of the hardness index of a single kernel measured on the Single Kernel Characterisation System (SKCS) is difficult to determine, since the experiment cannot be repeated (Nielsen et al., 2003). Averaging over several kernels provided better models for deter-

mination of bulk hardness (Maghirang and Dowell, 2003, Nielsen et al., 2003).

Clearly the accuracy of both near-infrared characterisation and the reference method must be good in order to establish prediction models on the level of single kernels. In a TOS perspective the single kernel constitutes the lot to be characterised and potentially sorted accordingly. By studying Fig. 3.2 it is clear that each individual kernel is quite heterogeneous themselves. Hence, assessing the entire chemistry non-destructively of a single kernel may not be straight forward. Sampling the kernel with NIT or NIR spectroscopy should ensure representative data of the entire kernel, the lot, in order to expect sufficiently robust models of the properties of interest. In the reference method, e.g. protein or hardness, the entire kernel (lot) is destroyed and the experiment cannot be repeated. Hence, the error in this destructive analysis is purely due to analytical error as no sampling is taking place. In Fig. 3 in Paper I, the error with respect to sample presentation of single kernels in the NIT beam was investigated. It showed that the root mean squared error of cross validation (*RMSECV*) in PLS protein prediction model was highly influenced by kernel presentation in the spectrometer. Using a single recording, pointing the furrow of the grain down, while the beam enters from the top should be preferred to other directions or random positioning. *RMSECV* was in the order of 0.4 to 0.6 with the furrow down, while in the order of 0.5 to 0.7 using random positioning, when considering ±2×standard deviation of the *RMSECV*. This was similar to the performances of Delwiche (1998) in reflectance mode with the furrow pointing away from the beam. By averaging several recording better performances was obtained. While covering several directions more representative recordings should ideally be obtained for minimization of kernel shape effects (Abe et al., 1996 cf. Pasikatan and Dowell, 2004). However, the results from Paper I showed that averaging three recordings of the kernels in the same position with the furrow down, ensuring representativity by re-positioning it between each recording, completely outperformed averaging over different directions. *RMSECV* went to 0.36 %dm protein which compared to the standard deviation of the reference method of 0.16% was considered good. The reason for this apparently biased method being the best might be due to unfortunate scattering patterns when recordings were averaged from different directions.

In the TriQ sorter used in Paper II and III the kernels are measured only once and there are no controls of the kernel orientations (Paper I) in the measurement area, however this is of no practical importance as this would complicate and slow down the sorting mechanism significantly. The twenty-four scans for each kernel used by Dowell et al. (2006a) in their sorting devise seem superfluous considering the high accuracy of the NIR technology and the corresponding lower accuracy of most reference methods. The increased time consumption on the NIR recording is probably not well spent when considering the capacity needed in such devises for practical and commercial use. This would be counteracting the purpose of the sorter for fast quality sorting of bulk materials. The performance of single recordings is sufficiently good relative to the variation found in different bulk lots.

The potential of utilising SKNIR to sort according to an internal quality trait has now been acknowledged. As protein content as well as its quality is very important for loaf volume (Bushuk et al., 1969), sorting with respect to protein content may increase the value (Dowell et al., 2006a, Pasikatan and Dowell, 2004). Pasikatan and Dowell used a colour sorter equipped with a dual wavelength filter (920/1660nm) for separation high protein content kernels from low protein content kernels. However the method was not very effective and several passes through the sorter were necessary. Each pass changed the resulting protein content by 0.1%-points. Dowell et al. (2006a) presented a system, which were able to record full spectra 950-1650 nm of individual kernels and successfully sort them into four different bins according protein and hardness calibration. The capacity of the system was 30 kernels/min. The system share some of the features of the TriQ sorter (Löfqvist and Nielsen, 2003) however the capacity was way too low for bulk sorting. Results from the laboratory scale TriQ bulk wheat sorter with a capacity of sorting 2 kg/hour (Löfqvist and Nielsen, 2003) utilised in Paper II and III were first reported in Nielsen (2002) showing a distinct sorting of grains with respect to protein content into three fractions. Raw wheat with average protein content of 12.3% was sorted into protein fraction <11%, 11-13% and >13%. Due to the considerable variation in the raw material the average protein fraction concentrations were 10.2%, 12.0% and 14.4% respectively with yields of 26%, 38% and 36%. Thus the high protein fraction was improved by 2.1%-points compared to the starting material. Munck (2008) report Farinograph and Extensiograph results from fractionation with the TriQ sorter based on a calibration to bread volume. In three fractions of 35%, 45% and

20% yield the dough stability time was 1.7, 5.5 and 8.4 min, water uptake was 53.1%, 56.7% and 59.7%, dough elasticity height 100, 129 and 146 and wet gluten content 17.4%, 22,5% and 27.6% respectively corresponding to low, medium and high quality.

Other commodities may be sorted in a similar way. Armstrong (2006) reports on an SKNIR system for measuring individual corn and soybean moisture and soybean protein content as the kernels slide down through a measurement area. With a capacity of 10 kernels/s it is to be fitted with automatic feeding and sorting mechanisms currently being developed. Ritteron et al. (2004) suggest a SKNIR system for sorting brown rice with regard to moisture and protein. Sorting *Fusarium* infested grains for reduction of deoxynivalenol (DON) content in bulk wheat was has also been reported (Delwiche et al., 2005)

In Paper II the effects on grains, flours and breads of sorting four different mixture lots from two different growth years and two different locations/cultivars were presented. The sorting was based on a two component PCA model of SKNIR from lots with known baking quality. The lots were sorted into three equally sized fractions according to their PCA scores (Löfqvist and Nielsen, 2003). In total fifty-two parameters were measured to characterise the wheat lots from grain to brain. It was shown that the major variance in the material tested was due to climate (i.e. growth year) and location/cultivar with no possibility to separate the effects of location and cultivar as they were confounded in the experimental plan. The analysis was made using both PCA and ANOVA which provided complimentary information and showed that the redundant information on the three levels, grain, flour and bread was coherent and thus validated each other. In Fig. 3.3 parameters reflecting significant effects regarding sorting with respect to fractionation by SKNIR sorting are presented. Panels A to E are parameters determined at grain level on a Foss Infratec 1241™ which shows a pronounced effect on protein levels (A), bulk grain density (B), Zeleny sedimentation value (C), starch (D) and wet gluten content (E). Panels F to L shows similar results on the flour level on protein content (F), wet gluten (G and L), falling number, Farinograph water absorption (I), Softening (J) and gelatinisation temperature (K).

Although significant results were found on parameters measured on the bread level (Fig. 3.3 panels M to P), the diversification results were not similarly convincing. The significance of final bread mass (M) and sensory perceived dryness (P) can easily be explained by the fundamental difference in water absorption (I), while the TPA chewiness of fresh bread (N) and TPA cohesiveness of thawed bread (O) were not providing satisfactory trends with regards to fractionation.

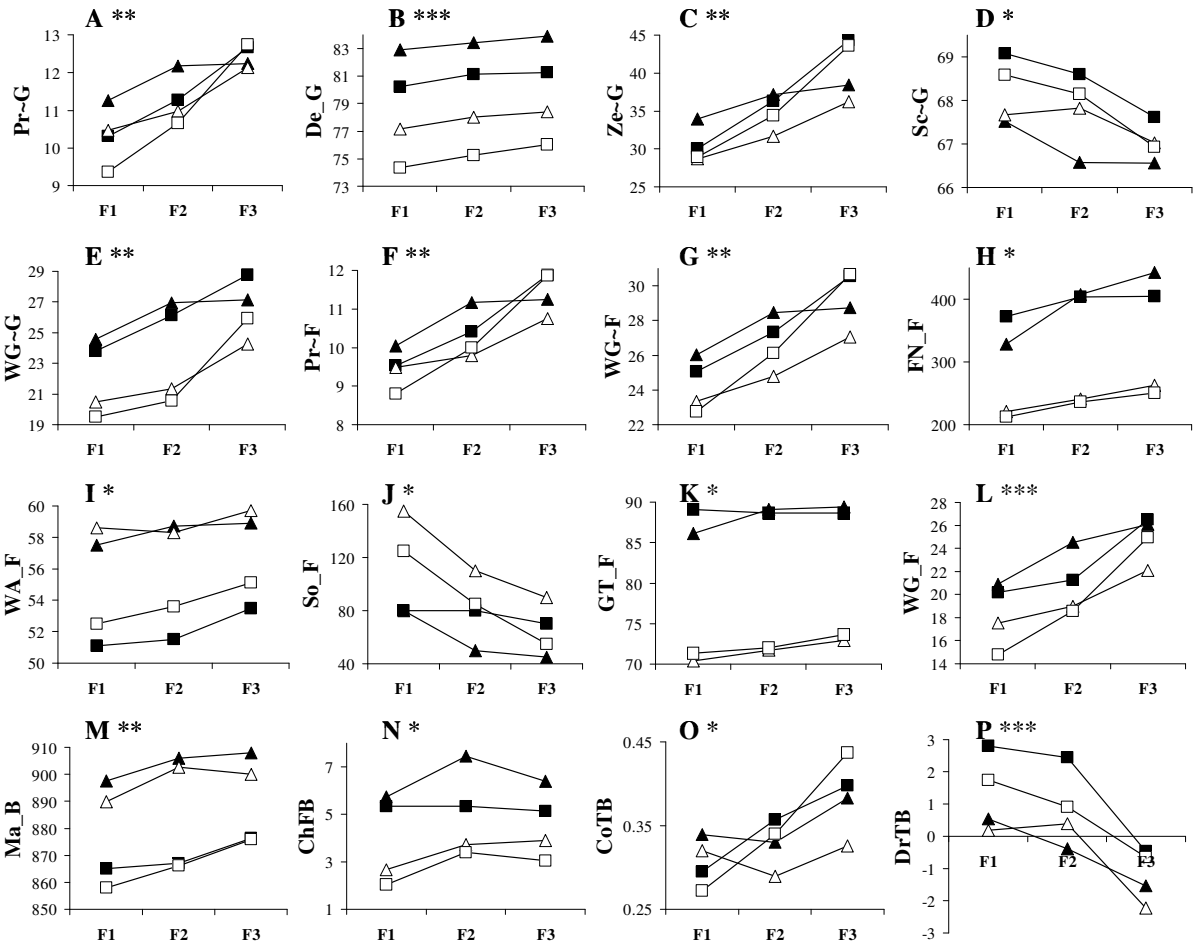


Figure 3.3: Raw data plots of significant parameters from the fractionation treatments (Table 3), 03AaVi (■), 03KiCo (▲),04AaVi (□) and 04KiCo (△). Se Table 2 in Paper II for specifications of parameters presented.

While the TriQ SKNIR fractionation clearly affected the protein level and quality as well as the amylase activity, the results regarding baking quality were not clear (Paper II). Significantly larger protein content and quality would normally lead to higher bread volume (Bushuk et al., 1969), thus fraction 1 to 3 should provide increasing bread volumes relative to year and location/cultivar. This was however not the case and the baking results in-

cluding the subsequent instrumental and sensory evaluation does not appear to be coherent with grain and flour results. The possibility of sample confusion with respect to flours used for the baking tests cannot be completely ruled out; however the baking tests using automatic baking machines may be the main source of unfortunate results (Chap. 3.2.5).

For conceptual validation near-infrared transmission and infrared reflection spectra of the grains from fractions 1 to 3 for each of the years and location are presented in Fig. 3.4 from 850 nm to 1048 nm. A clear qualitative pattern is observed from the raw spectra. The overall absorption levels are highest in fraction 1 and falling through to fraction 3 in each group. This indicates that light absorption is decreasing with increasing fraction number. This might be due to the particle size distribution of the sample, thus small and rough particles leads to an effectively longer path length and loss of light which results in larger absorption. The NIT spectra of the grain lots in Fig. 3.4 thus confirm that the TriQ sorting actually physically as well as chemically diversify the properties of the grain fractions in comparison to the starting material as also seen in Fig. 3.3B in which the bulk grain density is increased with increasing fraction number. The unsupervised TriQ sorting calibration in the 1100 to 1700 nm band was made on data without prior scatter correction, thus a vast part of the quality information is found in the first component of the raw data.

In Fig. 3.5 multiplicative scatter corrected (MSC) IR spectra from 1700 to 1450 cm^{-1} of the flours of fractions 1 to 3 are presented similarly for each group. The Amide I band at 1650 cm^{-1} corresponding to the C-N and C=O stretching vibration and the Amide II band at 1540 cm^{-1} corresponding to the N-H bending and C-N stretching vibrations clearly confirms the previous findings above in that the protein levels are increasing in fraction 1 to 3. There are apparently no distinct differences in the shape of spectra of either the fraction or the group.

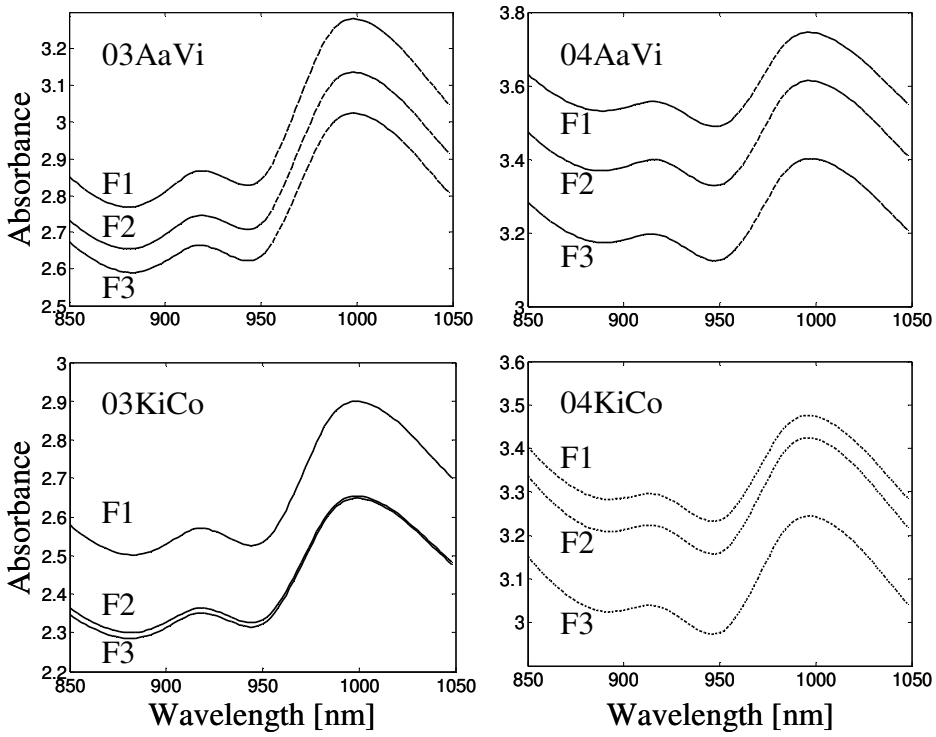


Figure 3.4: Raw NIT spectra from 850 to 1048 nm of fractions 1 to 3 grains in each year and location/cultivar. Increasing absorption is found with decreasing fraction number.

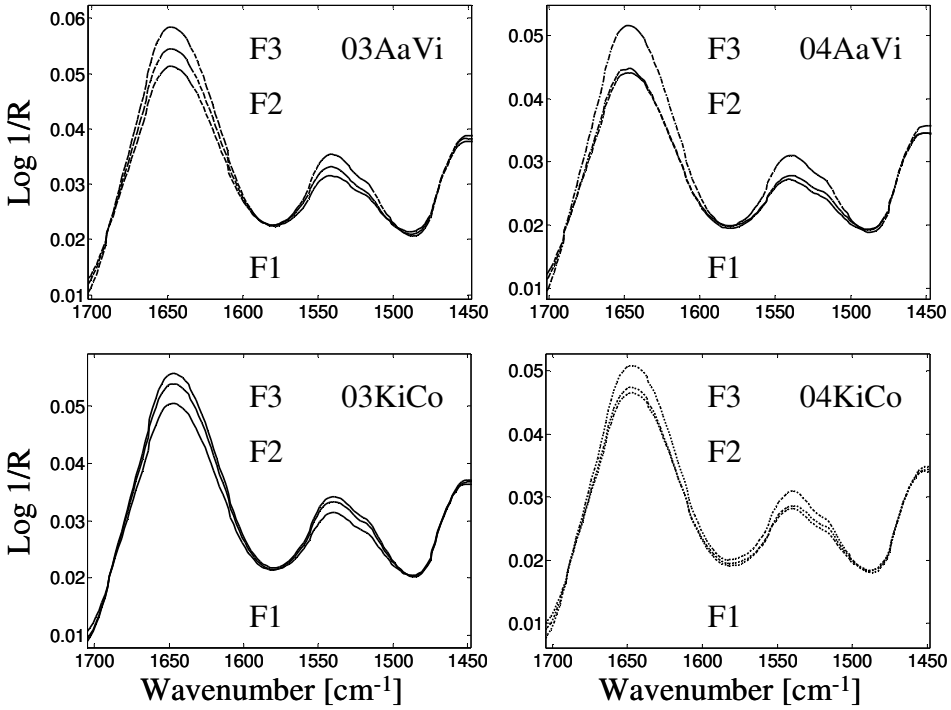


Figure 3.5: MSC treated IR spectra from 1700 to 1450 cm^{-1} of fractions 1 to 3 flours in each year and location/cultivar. The amid I and II bands at 1650 cm^{-1} and 1540 cm^{-1} respectively are clearly more intense with increasing fraction number.

3.2.5. Test baking using small automatic home-bakery

In order to evaluate the end effect of bulk material diversification on end product quality, breads had to be baked in a uniform and reproducible way. Since a trained test baker was not at hand, the use of automatic household baking machines to produce fresh bread was utilised. Many test laboratories have a battery of this type of machines available for rapid screening and testing. Several reports on their general applicability are available (Grausgruber et al., 2001, Hansen and Hansen, 1992, 1993, Peltonen and Salovaara, 1991, Zwingelberg and Brümmer, 1990). Although the machines are generally considered reproducible they may also mask some of the effects of varying flour quality, as when some machines bake relatively large breads compared to detached breads of flour with low protein and low Zeleny sedimentation volume (Zwingelberg and Brümmer, 1990). The home-bakery used by Grausgruber et al. (2001) produced reliable discriminating results only when mixing was conducted outside the machine – thus not being able to utilise the automatic feature. While using an automatic baking program similar to the one used in Paper II and III, they were unable to discriminate between flours classified in nine quality groups (QG) covering the entire quality range from feed wheats to high quality bread wheats with protein contents from 12.6% to 15.6%. In Paper II and III four Panasonic SD-253 units were used to assess the baking quality using a standard bread recipe. The machines did produce relatively uniform breads with a relative standard deviation on volume below 3%. This was sufficient to discriminate effects of growth year and location/cultivar for a range of bread parameters such as mass, volume, hardness and elasticity. However results from Paper II and III showed that loaf volume was not directly correlated to protein and gluten as expected (Bushuk et al., 1969, Peltonen and Salovaara, 1991) rather to falling number, gelatinisation temperature and gelatinisation maximum (See below).

Bushuk et al. (1969) showed that baking test using the same mixing time for different cultivars and different protein contents may mask the true baking potential of the lots involved. Weak cultivars may be over-mixed, while strong wheats may be under-mixed which both result in smaller loaf volumes than potentially possible. Likewise within a cultivar with different protein contents, the mixing time for the gluten to develop should be increased with decreasing protein content to reveal its true baking potential. While at least some automatic home bakeries have a tendency to favour

weak doughs due to unfortunate mixing conditions they may still serve as semiautomatic devices by mixing the dough outside the machine (Grausgruber et al., 2001, Zwingelberg and Brümmer, 1990) in controlled mixing conditions e.g. in a Farinograph (Peltonen and Salovaara, 1991). In the Panasonic SD-253 this would involve removing the mixing blade and covering the hole and shaft in the bottom of the pan by metal foil.

3.2.6. SKNIR sorting in the future

A commercially available laboratory scale single seed sorter (Luminar 3076 "Seed Meister" NIR Analyser, Brimrose, Maryland, USA) has been available since 1996 for plant breeding purposes (Nielsen, 2002). The breeding perspective of sorting is indeed emphasised by others (Dowell et al., 2006a, Munck, 2008, Nielsen et al. 2003, Osborne, 2006). Munck (2008) foresee great potentials of bulk sorting if it can be made economically viable by increased market value of the product, and for breeding purposes. Comprehensive studies of genetic and environmental effects on sorting should however substantiate the benefit of sorting for breeding purposes.

With regards to how to physically overcome individual kernel characterisation, classification and subsequent placing in the appropriate receiving bins with respect to the vast amounts of kernels in bulk material is a matter of inventive engineering. Löfqvist and Nielsen (2004) solved this in a scalable invention which potentially is capable of sorting several tons of grains or other particulate materials per hour.

3.3. Predictions of flour and bread quality

"The ability to predict reflects a high degree of process understanding" (US-FDA, 2004a). With the vast amount of data available from grain, flour and bread level the potential of predicting end product quality is certainly present. Especially near-infrared spectroscopy is generally used for routine quality assessments in breeding as well as in the cereal industry (Osborne, 2006). In Paper III, the aim was to screen various fast spectroscopic methods for predicting manifest parameters determining flour quality and in turn use the same method to predict bread quality as determined by test baking in the automatic baking machines.

3.3.1. Predictions of flour functionality

Near-infrared spectroscopy has long been used for rapid testing of flour in relation to it functionality in bread making. Osborne et al. (1982) described successful multiple linear regression (MLR) calibration to moisture, protein, particle size, colour and damaged starch using a filter instrument with ten filters in the range from 540 to 2310 nm. While Delwiche and Weaver (1994) were able to predict water absorption reasonably well, they were unable to predict mixing time and mixing tolerance using NIR. Similar results were obtained by Dowell et al. (2006b) who were unable to establish good PLS prediction models to most Mixograph, Farinograph and Alveograph parameters apart from Mixigraph absorption. Mirablés (2003, 2004) on the other hand showed that NIT of flour is a good method for predicting both Farinograph and Alveograph parameters. The results of Paper III regarding prediction of water absorption from with NIR is agreement with Mirablés (2004), however the results for stability and softening did not compare. Delwiche et al. (1998) successfully established NIR models for glutenin, gliadin, Zeleny sedimentation and Mixograph peak resistance. In a comprehensive study Dowell et al. (2006b) used four different NIR instruments for predicting 46 and 47 grain, milling, flour, dough, and breadmaking quality parameters of 100 hard red winter and 98 hard red spring wheat samples respectively covering the quality ranges of U.S. commercial wheat. Although high correlations were found for a number of parameters it was concluded that this was primarily caused by the strong correlation of these parameters to protein content (Dowell et al., 2006b, Mirablés, 2004). Apparently successful modelling is dependent on the diversity of the wheat material assessed (Delwiche et al., 1998).

In Paper III it was shown that near-infrared spectroscopy outperforms infrared of flour and nuclear magnetic resonance (NMR) relaxation profiling of the baking process for prediction of a large number of flour quality parameters simultaneously. Protein content, wet gluten content and Farinograph water absorption were well modelled flour parameters with 99.0%, 95.8% and 95.7% of the variation explained using NIR. The α-amylase activity was similarly good modelled; 92.2%, 94.8% and 89.1% of the variances in gelatinisation temperature, gelatinisation maximum and falling number respectively were explained. The Farinograph parameters, development time, stability and softening were less well determined as was gluten index and sedimentation volume.

3.3.2. Predictions of bread quality

It is generally accepted, that protein content and protein quality are closely related but separate factors for determination of breadmaking potential in wheat. Roughly speaking, the quality of the protein is primarily determined by genetics, while the contents are determined by agronomical conditions; soil, fertilisers, climate and pests (Bushuk et al., 1969). The gluten proteins witch constitute 80 to 85% of the total wheat protein are responsible for the visco-elastic properties of dough. The monomeric gliadin fraction provides viscous properties, while the polymeric glutenin fraction is elastic and stabilised by inter- and intra-chain disulfide bonds. While the complex discussion regarding gluten structure, composition and interaction with other wheat constituents such as lipids, arabinoxylans and non-gluten proteins is omitted in this thesis, the functionality measured by Zeleny sedimentation, gluten content, gluten index and rheology in the Farinograph are indeed indirect measures of the inherent gluten quantity and quality. Other constituents in wheat play a role regarding the functionality of wheat, but it is primarily the quantity and quality of the gluten proteins which govern the bread making functionality of wheat (Veraverbeke and Delcour, 2002). However, Greybosch et al. (1993) found a number of biochemical components (lipid, protein and pentosan fractions) were necessary for determination of dough handling and loaf characteristics.

Being able to predict protein content as well as the quality, chemical as well as rheological, and α-amylase activity may provide a fast one-step procedurere of great potential in the milling industry and in breeding programmes (Osborne, 1984). Devaux et al. (1986) was able to discriminate and classify common wheat varieties and three classes of wheat based on their recognised baking quality using NIR and a combination of principal component analysis (PCA) and multiple discriminant analysis (MDA). Direct prediction by Delwiche and Weaver (1994) of bread loaf height, internal grain appearance and overall bake score by NIR and PLS was unsuccessful. Dowell et al. (2006b) were able to predict loaf volume, but not crumb grain score

Lots of the same cultivar but with varying protein contents usually result in bread quality with respect to loaf volume which is relatively linear correlated to the protein content over a wide range (9-16%) using the same baking test. The bread quality of different cultivars with the same protein content

on the other hand is determined by a number of factors related to the specific cultivar (Bushuk et al., 1969).

In the PLS and multiblock PLS modelling in Paper III it was shown that the functionality parameters of flour had an unsurpassed ability to predict end bread quality. However adding data blocks of NIT on grain and flour improved the predictive performance of several parameters. Adding the NMR-relaxation profiles improved predictions further for a few not well determined texture profile attributes, namely cohesiveness, springiness and resilience. Especially bread mass, volume and density prediction models could be established emphasising the feasibility of such endeavours. However PLS models for TPA hardness, chewiness and gumminess and associated sensory perceived attributes; hardness, chewiness and fracturability were also satisfactory considering the type of data which can be inherently noisy. Details may be viewed in Paper III.

3.3.3. Process understanding gained

The PLS is a fantastic tool for prediction of various chemical and physical properties from data containing several interferences which need modelling along with the properties of interest. On the same time the graphical approach of the chemometric method provides an opportunity to look inside the calibration and pinpoint the parameters or variables which are important and covarying with the parameters we wish to predict. Specifically the regression coefficients may be used directly as a diagnostic tool (Paper III, Fig. 7).

It was shown that a number of instrumental and sensory texture attributes had virtually the same regression coefficients as the bread volume determined by the twelve flour functionality parameters only. Thus many of the attributes were responding to the same under lying phenomenon. In Table 3.1 the correlation coefficients, r, between flour quality and bread quality is shown for further elucidation. Correlation coefficients larger than 0.7 have been emphasised with bold numbers for easy pattern recognition. The volume of bread in this analysis is governed primarily by falling number (FN_F) and the gelatinisation temperature, which were both negatively correlated (r = -0.75) to volume. The TPA hardness, gumminess and chewiness as well as sensory perceived fracturability, hardness and chewiness are all well correlated to bread volume ($0.87 < |r| < 0.91$). Two major lessons can be

derived. The good part is that these instrumental and sensory texture parameters are highly correlated, which may be confirmed directly in Table 3.1. Future measurements of hardness and chewiness may thus be performed instrumentally only. The second part is that volume naturally governs these texture parameters. Voluminous bread crumb must be softer than more compact bread – simply due to more empty space being compressed in the tests, be it instrumental or sensory.

Similarly sensory perceived elasticity and dryness had similar correlation coefficients as those for bread mass. The bread mass is governed by the water absorption as shown in Fig. 7B in Paper III and so is elasticity and dryness. This may also be confirmed in Table 3.1. The water absorption was primarily diversified by location/cultivar, i.e. lots from Aarlev/Vinjett had water absorption of 51 to 55%, while Kiel/Vinjett and the remaining had water absorption of 57.5 to 62%.

A close look a Table 3.1 also shows that bread volume is not correlated to protein content and Zeleny sedimentation volume at all. These parameters are usually correlated to bread volume (Bushuk et al., 1969). This information is pointing to problems with the baking procedure as already discussed above (Chap. 3.2.5). The large variation in falling number associated with the growth year (Paper II) and the unfortunate preference for weak dough in the baking machines (Grausgruber et al., 2001) is determining the outcome of this experiment. Apart from being slightly disappointing in light of the possibility of bringing forward the benefits of SKNIR sorting on the end bread quality; this is a perfect lesson in the power of gaining process insight via multivariate analysis of the entire process from grain to brain. One strength in this study which contained relatively few experimental units was that the entire history of each individual lot was known, making full traceability possible regarding the original causes of the important variabilities identified (α-amylase activity and water absorption). Better prediction models may be obtained by ensuring even more diverse material. A more optimal process would involve a modification og the baking process, e.g. by mixing outside the baking machines.

Table 3.1: Correlation coefficients of flour quality and bread quality parameters used in Paper III. Coefficient larger than 0.7 is emphasised with **bold** types.

	r	De_G	Pr-F	FN_F	WA_F	DT_F	St_F	So_F	Ze_F	GT_F	GM_F	WG_F	GL_F	Ma_B	WL_B	Vo_B	De_B	H1FB	H2FB	AFFB	CoFB	SpFB	GuFB	ChFB	ReFB	AdFB	YeTB	ElTB	FrTB	HaTB	DrTB	CwTB	AhTB	AfTB
Bread Sensory	AfTB	-0.02	0.05	0.33	**-0.72**	0.37	0.34	0.10	-0.48	0.33	0.65	0.04	**-0.72**	-0.59	0.11	-0.04	-0.08	-0.09	-0.01	-0.39	0.33	0.24	-0.03	-0.01	0.38	-0.38	0.45	0.40	0.05	0.02	0.59	-0.08	-0.11	1.00
	AhTB	0.39	0.05	0.42	0.53	-0.01	-0.05	-0.16	-0.20	0.41	0.29	0.26	-0.12	0.65	-0.66	-0.56	0.60	0.43	0.42	0.57	-0.41	-0.35	0.42	0.40	-0.34	0.54	-0.56	**-0.71**	-0.61	0.58	-0.47	0.69	1.00	
	CwTB	0.65	0.12	**0.74**	0.35	0.07	0.23	-0.56	-0.04	**0.75**	0.55	0.39	-0.20	0.56	**-0.79**	**-0.89**	**0.90**	**0.81**	**0.81**	**0.74**	-0.40	-0.40	**0.82**	**0.81**	-0.26	0.60	-0.33	-0.43	**-0.88**	**0.92**	-0.21	1.00		
	DrTB	-0.08	-0.32	-0.03	**-0.79**	-0.02	0.08	0.28	-0.43	-0.03	0.30	-0.39	-0.23	**-0.73**	0.31	0.06	-0.17	-0.14	-0.10	-0.41	0.24	0.14	-0.10	-0.09	0.21	-0.45	0.70	**0.76**	0.23	-0.08	1.00			
	HaTB	0.68	0.18	**0.77**	0.21	0.05	0.36	-0.60	0.02	**0.74**	0.61	0.40	-0.21	0.44	-0.75	**-0.91**	**0.89**	**0.87**	**0.87**	0.67	-0.43	-0.37	**0.87**	**0.87**	-0.24	0.52	-0.17	-0.31	**-0.87**	1.00				
	FrTB	-0.61	-0.18	**-0.72**	-0.37	-0.12	-0.39	0.63	-0.12	**-0.74**	-0.53	-0.44	0.18	-0.57	**0.80**	**0.87**	**-0.89**	**-0.78**	**-0.79**	-0.65	0.56	0.25	**-0.80**	**-0.81**	0.17	-0.50	0.36	0.44	1.00					
	ElTB	-0.29	-0.04	-0.20	**-0.81**	0.20	0.13	0.12	-0.04	-0.22	0.06	-0.23	-0.15	**-0.84**	0.59	0.41	-0.50	-0.38	-0.33	-0.57	0.56	0.35	-0.33	-0.31	0.46	-0.53	**0.88**	1.00						
	YeTB	-0.30	-0.01	-0.12	**-0.86**	0.24	0.13	0.12	-0.08	-0.16	0.17	-0.21	-0.20	**-0.86**	0.56	0.30	-0.41	-0.28	-0.22	-0.51	0.50	0.35	-0.23	-0.21	0.49	-0.45	1.00							
Bread TPA	AdFB	0.35	0.12	0.33	0.51	-0.07	-0.08	-0.43	0.17	0.34	0.07	0.27	0.14	0.59	-0.53	-0.50	0.58	0.56	0.50	**0.95**	-0.55	-0.52	0.51	0.48	-0.56	1.00								
	ReFB	-0.16	0.23	0.06	-0.43	0.46	0.26	0.08	0.03	0.09	0.17	0.15	-0.43	-0.47	0.42	0.33	-0.37	-0.34	-0.25	-0.60	**0.92**	**0.87**	-0.24	-0.18	1.00									
	ChFB	**0.81**	0.35	**0.87**	0.21	0.24	0.57	**-0.73**	0.17	**0.84**	0.64	0.58	-0.30	0.44	**-0.79**	**-0.87**	**0.87**	**0.98**	**0.99**	0.67	-0.41	-0.33	**1.00**	1.00										
	GuFB	**0.80**	0.32	**0.86**	0.23	0.21	0.54	**-0.72**	0.16	**0.83**	0.62	0.56	-0.27	0.46	**-0.80**	**-0.88**	**0.89**	**0.99**	**1.00**	0.70	-0.46	-0.39	1.00											
	SpFB	-0.33	0.21	-0.16	-0.31	0.32	0.15	0.18	0.13	-0.14	-0.05	0.08	-0.27	-0.43	0.55	0.48	-0.50	-0.48	-0.40	-0.59	**0.84**	1.00												
	CoFB	-0.34	0.11	-0.14	-0.47	0.40	0.14	0.17	0.00	-0.09	0.01	-0.01	-0.28	-0.56	0.58	0.50	-0.54	-0.56	-0.47	-0.65	1.00													
	AFFB	0.53	0.11	0.46	0.56	-0.08	0.03	-0.52	0.18	0.47	0.17	0.33	0.10	0.68	**-0.70**	-0.68	**0.75**	**0.74**	0.69	1.00														
	H2FB	**0.80**	0.32	**0.86**	0.21	0.22	0.54	**-0.72**	0.15	**0.83**	0.64	0.56	-0.28	0.44	**-0.79**	**-0.88**	**0.88**	**0.99**	1.00															
	H1FB	**0.78**	0.28	**0.81**	0.28	0.14	0.48	**-0.70**	0.15	**0.78**	0.56	0.51	-0.20	0.50	**-0.81**	**-0.87**	**0.89**	1.00																
Br. Dim.	De_B	**0.76**	0.02	**0.73**	0.40	-0.05	0.25	-0.54	-0.04	**0.74**	0.53	0.33	-0.14	0.63	**-0.88**	**-0.98**	1.00																	
	Vo_B	**-0.74**	0.00	**-0.75**	-0.27	0.04	-0.26	0.51	0.10	**-0.75**	-0.60	-0.29	0.21	-0.51	**0.85**	1.00																		
	WL_B	**-0.75**	-0.22	-0.69	-0.51	-0.09	-0.33	0.55	-0.08	-0.66	-0.45	-0.49	0.22	**-0.74**	1.00																			
	Ma_B	0.50	0.10	0.25	**0.95**	-0.22	-0.06	-0.30	0.27	0.25	-0.14	0.33	0.23	1.00																				
Flour Functionality	GL_F	-0.37	-0.47	-0.64	0.39	-0.69	-0.56	0.23	0.14	-0.60	**-0.74**	-0.54	1.00																					
	WG_F	0.56	**0.92**	0.67	0.19	**0.72**	**0.76**	**-0.76**	0.57	0.61	0.43	1.00																						
	GM_F	0.55	0.25	**0.87**	-0.38	0.55	0.58	-0.43	-0.32	**0.86**	1.00																							
	GT_F	**0.74**	0.35	**0.96**	0.01	0.50	0.61	-0.64	-0.08	1.00																								
	Ze_F	0.04	0.69	0.00	0.31	0.23	0.40	-0.58	1.00																									
	So_F	-0.51	-0.66	-0.67	-0.12	-0.49	**-0.74**	1.00																										
	St_F	0.46	**0.73**	0.67	-0.24	0.68	1.00																											
	DT_F	0.28	**0.72**	0.52	-0.33	1.00																												
	WA_F	0.30	0.02	-0.01	1.00																													
	FN_F	**0.76**	0.43	1.00																														
	Pr-F	0.28	1.00																															
	De_G	1.00																																

4. Chemometric method development

Analysis of NMR relaxation data is not straight forward. For prediction purposes and exploratory studies using the entire relaxation curve as a spectrum apparently works well (Engelsen et al., 2001). The good correlations found with regard to texture and staling was the inspiration to use the entire baking profile for prediction of bread quality (Paper III). However the interpretation of loadings is erratic in this type of investigation as they are difficult to relate to the physical compartmentalisation of water and fat in the sample. Instead the fundamental exponential relationship may be utilised for both quantification and interpretation.

$$M(t) = \sum_{i=1}^{I} M_{2i} e^{\frac{-t}{T_{2i}}} + f(t) \qquad \text{(Eq. 4.1)}$$

Thus the individual decay curves (Fig. 2.6) in the CPMG-baking arise physically from exponential decaying magnetisation. As the decay rates depend upon the local physical environments in the sample the measured decay, $M(t)$, is a sum of exponentials, each with a specific relaxation constant (T_{2i}) and amplitude (M_{2i}). In the window of observation only relatively mobile protons from water and fat can be observed, while protons bound in starch and protein decay so fast that no signal is recorded. Thus more rotational freedom results in higher T_2-values. With this knowledge it is possible to fit Eq. 4.1 to a number of exponentials, either discretely or as a distribution and assess the various pools of T_{2i}-values and their relative abundance, M_{2i}. The residual, $f(t)$, may be used to asses the goodness of the fit. The CPMG decay presented in Fig. 4.1 were recorded on a 23.2 MHz Maran Benchtop pulsed NMR Analyser, Resonance Instruments (Whitney, UK) using echo time, τ = 100 μs and 8190 echoes. Only even numbered echoes were used. Details may be found in Paper III.

4.1. Exponential fitting of spin-spin relaxation in the baking process

The discrete fit and the continuous fit are used interchangeably. They both have their advantages and disadvantages, but used together they may actu-

ally complement each other for better or at least more balanced interpretation. The distributed fit is thought to provide results, which are physically meaningful, as the idea of rigid discrete conditions in the sample seems unlikely. However the algorithms for fitting such problems suffer from being overwhelmingly labile (Butler et al., 1981). The mathematical problem is as such ill defined by the objective goal of estimating an endless number of parameters to a large number of correlated data. This problem is solved by a smoothing factor, which is a trade of between mathematically optimal solutions and a desired amount of smoothness. Observe three exponential decays in Fig. 4.1 recorded at the beginning of rising at 34°C, at 71°C during baking and at 34°C after cool down.

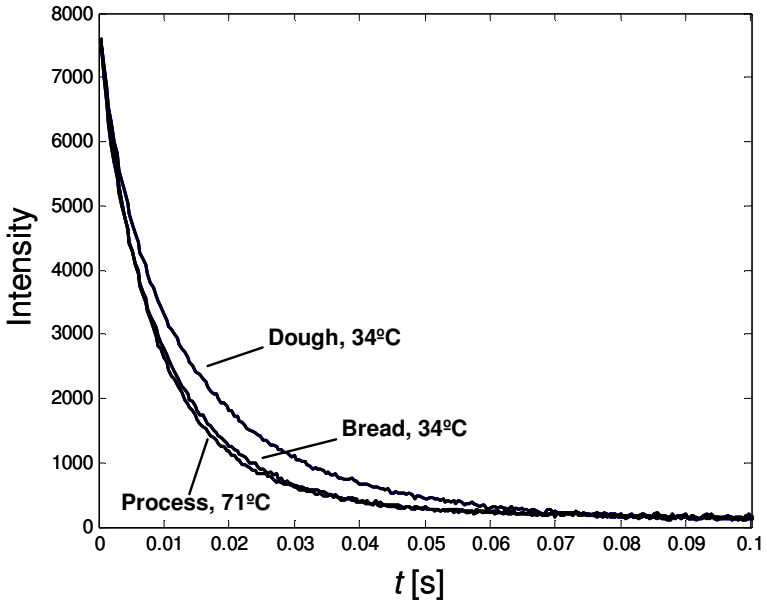


Figure 4.1: Max-normalised relaxation spectra of sample 04KiCapo at three stages in the baking process, as dough just after mixing at 34°C, in the baking process, at 71°C and as bread after cooling to 34°C. Different strategies for multi-exponential fitting are presented in Fig. 4.2 and Fig. 4.3.

In Fig 4.2 the goodness of fit was assessed by inspecting the root mean squared errors (*RMSE*) of $f(t)$ of A: distributed solutions with respect to smoothing weights and B: discrete solutions with respect to number of discrete components. Note here that *RMSE* of the optimal distributed and discrete models are equivalent. The apparently optimal weights are somewhat varying depending on the decay curve for the distributed solution. Experience with the algorithm shows that choosing the same smoothing parameter for all samples should be preferred when more samples are to be compared,

thus weight = 0.1 was chosen (Fig. 4.2A). In the discrete case, the complexity apparently increases from three to four after cooling (Fig. 4.2B).

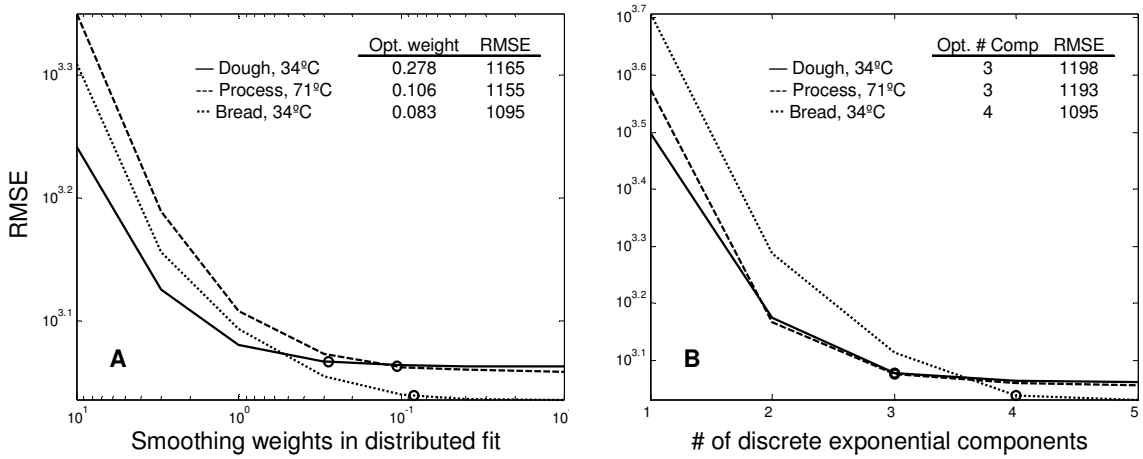


Figure 4.2: Root mean squared errors (*RMSE*) of exponential models of sample 04KiCapo presented in Fig. 4.1 and 4.3 **A**: *RMSE* as a function of smoothing weights in distributed fits at three points in the baking process. **B**: *RMSE* as a function of number of discrete exponential components at three points in the baking process.

In Fig. 4.3 first column the exponential fits from Fig. 4.2A using the algorithm of Butler et al. (1981) is shown (Fig. 4.3A, C and E). By varying the weight from 0.01 to 10 in the algorithm an increasingly smoothed result is obtained. The problem here is that all the results are solutions to Eq. 4.1. And since the results are determined by data as well as more or less subjective assessment of appropriate smoothness how may we trust the interpretation of the result as a truly objective analysis. It may be very difficult to determine whether a certain peak is really a physical component or rather an unfortunate result of regularisation.

In the 2nd column of Fig. 4.3 the distributed solution weight = 0.1 is superimposed with distributed fits assuming either two, three or four components (Fig. 4.3B, D and E). The distributed fit shows an intense peak around 10 ms throughout the dough (34°C), process (71°C) and bread (34°C). In the dough a tail towards the fast T_2 values around 1 ms is observed. This tail has budded off at 71°C with a distribution below 1 ms, while in bread, it has moved closer to the major peak again. A small peak above 100 ms is seen throughout, while a very slow component is present only in the bread – probably free water condensed from the vapour in the closed vial. The apparent assessment is a quite stable system with three major components and a forth showing up after cooling. However the very fast components may be due to

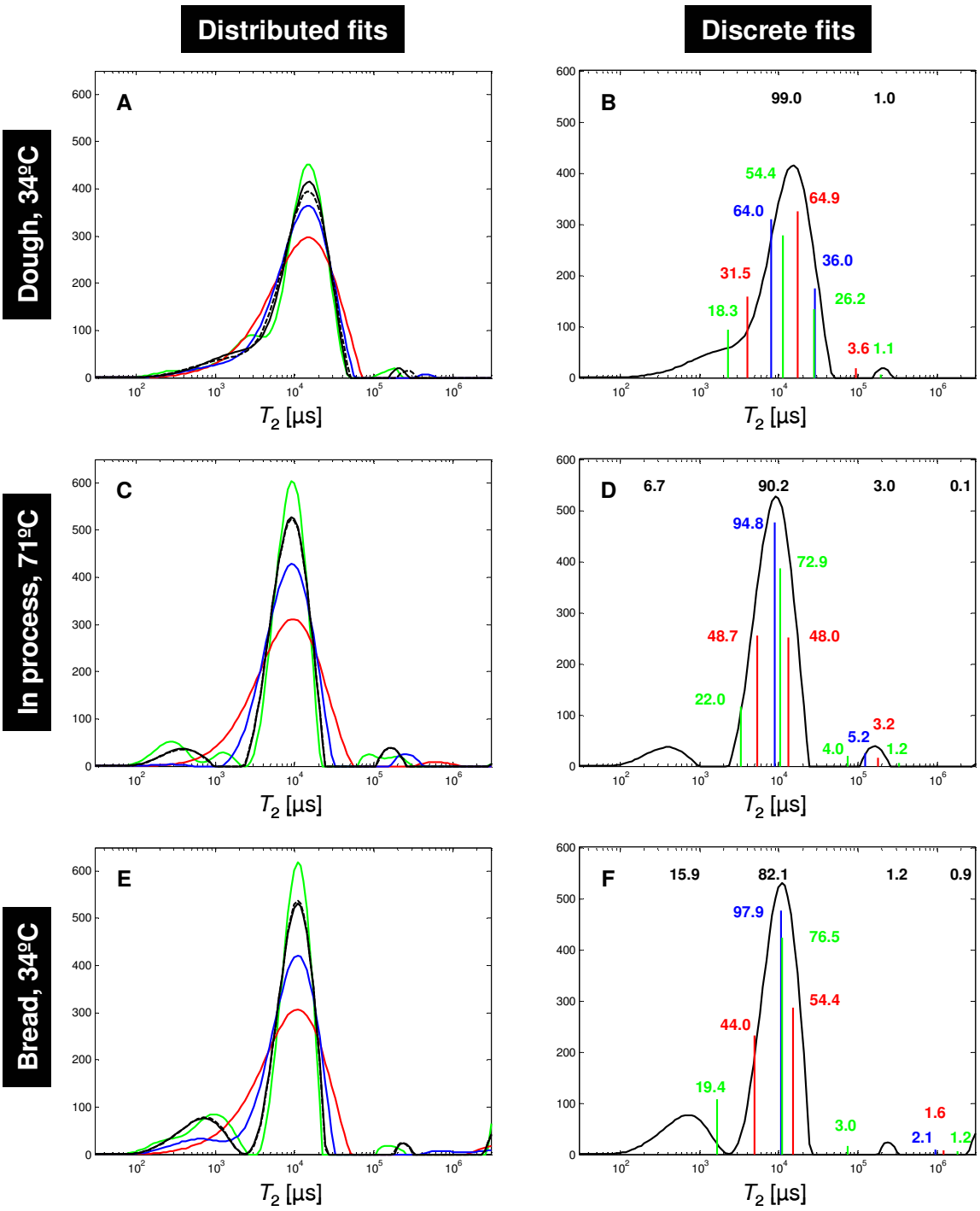


Figure 4.3: Multi-exponential fits of data presented in Fig. 4.1. First column display different distributed results depending on regularisation smoothing weights used in the Butler, Reeds and Dawson (1981) algorithm at three different stages (**A**, **C** and **E**) of the baking process using different weights: — weight = 0.01, — weight = 0.1, --- weight = optimised, — weight = 1, — weight = 10. Second column display three discrete fits of the same three stages (**B**, **D** and **F**) using two, three and four exponential components. Numbers refer to the relative abundance of the T_2 components. The curves included for reference are identical to the black curve (weight = 0.1) in the first column.

an over fit. With a distance of 400 μs between each measurements it should be clear that all results below 1 ms is based on a very few points since such components are totally decayed within a couple of ms.

The superimposed discrete fits are somewhat showing a similar pattern (Fig. 4.3B, D and F). Regardless of the number of components chosen, the bulk peak around 10 ms is clearly modelled by one or two discrete peaks. The small peak above 100 ms is also modelled by the three or four component model in dough, all models in process as well as in bread. In the bread however the two and three component models are using only one component to model both the above 100 ms and the free water component, while this is handled better by the four component model. The fast component is handled in dough and bread by the three and four component models respectively, while none of the discrete fits acknowledge the fast component below 1 ms in the process (Fig. 4.3D). By assuming a three component discrete model throughout rising and baking and a four component model after cooling the results are comparable to Engelsen et al. (2001).

For interpretation, determination of the number of components is crucial, whether it is in a distributed way or in a discrete way. Using jack-knife for determination of the number of components in individual transverse relaxation spectra as suggested by Pedersen et al. (2000) is not at stable solution to the problem. The results are also dependent on the number and size of the segments and how they are divided (123123123, 111222333 or randomly) all of which is decided by the investigator.

4.1.1 Suggestion for robust determination of discrete exponential model complexity.

An alternative approach based on the stability of the estimated T_2 times and concentrations, M_2 from a series recording of similar samples is presented below. The hypothesis is that in a stable mathematical solution the variation between the estimated parameters are smaller than in an unstable solution. In Fig. 4.4 two different wheat lots have been baked by the CPMG procedure, each in four replicates. Discrete exponential models have been calculated using one to five components for all eight experiments and all thirty time-points in the baking process. Thus at each time-point the root mean squares standard deviation (*RMSSD*) has been calculated for $\log(T_{2i})$ and M_{2i} values estimated from each sample and averaged over the components in

each solution and over the two different wheat lots. Thus Fig. 4.4 represents the stability of solutions with respect to number of components and time point in the baking process of $\log(T_{2i})$ and M_{2i}.

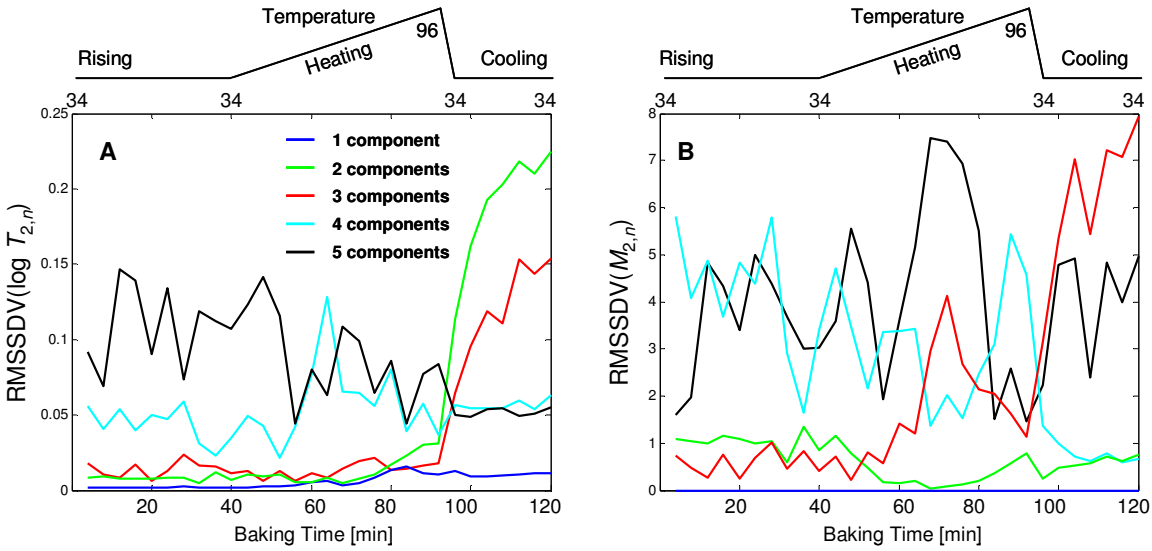


Figure 4.4: Root mean squared standard deviation (*RMSSD*) of discrete exponential solutions in the baking process with respect to **A**: $\log(T_{2i})$ and **B**: M_{2i}.

In Fig. 4.4 the one component model should be disregarded, as this is *a priory* regarded unlikely despite its mathematical advantage. With regard to $\log(T_{2i})$ the stabilities of both the two and the three component model are superior ($RMSSD < 0.03$) to models with more components during rising and baking from 0 to 92 minutes (Fig. 4.4A). Immediately at the time of cooling the two and three component models fail, while both the four and five components models appear as candidates with similar low *RMSSD* values around 0.05. With regard to M_{2i} the two and three component model seems reasonably stable through the rising time (0-56 min), with the three component model as the most stable (Fig. 4.4B). Shortly after heating starts (56 min) the two component model becomes superior until met by the four component model during the cooling (100 min). These results point to a new hypothesis regarding the exponential relaxation components during bread baking: During rising, a three component model is appropriate, while at approximately 50°C a two component model is adequate and after cooling down a four component model is taking over. This hypothesis is also supported by the distributed fit when the very fast component below 1 ms is regarded a consequence of over fit (Fig. 4.5).

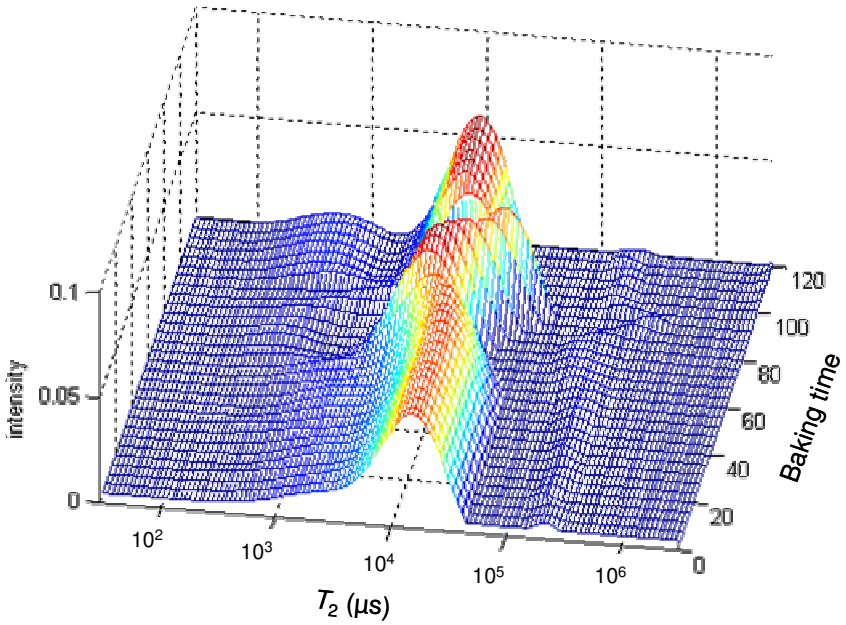


Figure 4.5: Distributed CPMG-fit of the entire baking process (Fig. 2.6). Apparently the dough and bread matrix is dominated by a bulk water ($T_2 \approx 10$ ms) with restricted rotational freedom and a minor component, probably associated with fat ($T_2 \approx 200$ ms).

By looking at the entire baking process using the distributed fit the development in T_2 components may be followed. The entire process is dominated by a single intense peak around 10 ms depending on the process stage. This peak shifts slightly to the left as the temperature increases. This implies a further restriction of rotational freedom of this water which could be due to the wetting of starch granules and the gelatinisation. With further increase in temperature, T_2 increases again, which may be due to temperature. The heat simply increases the rotational freedom of water. The generally low relaxation value of the bulk water is probably due to all water being more or less associated with surfaces within the matrix throughout the baking process. An even stronger bound component appears to be present ($T_2 \approx 1$ ms) at least during rising. Engelsen et al. (2001) associated this fast water component with protein, which looses affinity for water during baking. As was shown in Paper IV by varying the fat content, the small component at $T_2 \approx 200$ ms is probably due to fat.

Another promising discrete method is the SLICING method, which based on a number of samples with similar components provide a discrete solution

with identical T_{2i} values for all samples, while intensities of the components may vary (Engelsen and Bro, 2003, Pedersen et al., 2002b). The method however may also be used on single decay curves for individual assessment of T_{2i} values (Manetti et al., 2004).

4.2. Diffusion-relaxation correlation spectroscopy (DRCOSY)

In order to gain further information regarding the components of dough and bread the T_2 associated diffusion was assessed by 2D diffusion-correlation spectroscopy (DRCOSY). The idea is that different components exist which are characterised by different pairs of diffusions coefficients and spin-spin relaxation times (Paper IV). By adding an extra dimension, overlapping components may be resolved into separate components (Callaghan et al., 2003, Hubbard et al., 2005, Qiao et al., 2005). The multi-dimensional decay has been used in porous media such as plants (Qiao et al., 2005), food (Godefroy and Callaghan, 2003, Hubbard et al., 2005) and rock samples (Song et al., 2002).

As shown in Paper IV the 2D-Laplace inversion (Istratov and Vyvenko, 1999) often used to analyse this type of data has similar labile characteristics as the distributed fits shown above generating spurious. The determination of the smoothing factor is crucial for the interpretation and different levels of noise may be detrimental for the analysis of the data (Hürlimann et al., 2002, Song et al., 2002).

For this reason we suggested a new method, 2D PARAFAC-Laplace decomposition, for analysing DRCOSY data, and for that matter similar 2D NMR correlation data. By utilising the PARAFAC model (Bro, 1997, Harshman, 1970, Carroll and Chang, 1970) which has the ability to decompose multilinear data into unique components, a robust quantification and interpretation of the data can be established (Paper IV). The 2D diffusion-relaxation curves are tri-linear by definition as each point is a sum of intensities multiplied by two exponential functions (Paper IV, eq. 1). The PARAFAC components are themselves 2D exponentials and may be translated to the T_2-D domain by the Laplace inversion.

The method extracted two components of the dough data. Component one had a T_2 relaxation constant at 180 ms corresponding to the findings in 1D (above) with a diffusion constant $D = 3 \cdot 10^{-12}$ $m^2 s^{-1}$. This component varied

amplitude systematically with varying oil content of the samples. It was hypothysed that the oil being relatively rotationally free and on the same time restricted in diffusion as compared to the diffusion rate of free water (D = $2 \cdot 10^{-9}$ $m^2 s^{-1}$) was in the form of small vesicles trapped in water films in the gluten network (Belitz and Grosch, 1999).

The second component had two major peaks in the T_2-D domain both with a T_2 relaxation constant in the order of 10 ms corresponding to the findings in one dimension (above). The T_2 relaxation constant for bulk water is in the order of 100 ms (Song et al., 2002). The Laplace components had diffusion constants in the order of $D = 10^{-9}$ $m^2 s^{-1}$ and $D = 3 \cdot 10^{-13}$ $m^2 s^{-1}$. Thus the single bulk water peak domination the CPMG-baking process was in reality as sum of two components. In the present investigation they show up in the same PARAFAC component, hence they were correlated. It must be emphasised that the major peak although restricted in rotational freedom can diffuse with the rate of free water. The major part of the water is thus associated with the enormous surfaces in the dough matrix while being able diffuse unhindered around in the matrix. The other Laplace component is tightly bound in the matrix, probably associated with beta-glucan which can absorb large amounts of water compared to protein or it may be bound to damaged starch.

Paper IV showed a quantitative relation of the PARAFAC components to the known water and fat content and the method worked as a filter removing spurious peaks present in the raw data. With PARAFAC working on a set of 2D decay curves, the noise and small artefact occurring randomly in the data was removed when calculating the components.

Further work regarding these results is needed. Analysis of the baking process similar to that mentioned above in 1D may elucidate more of the water and fat dynamics available. By measuring diffusion-relaxation during baking and applying the PARAFAC-Laplace procedure at each temperature unique diffusion-relaxation spectra may be obtained. The dynamics of especially the water components should be interesting to follow. The PARAFAC-Laplace method however, works on a set at spectra in which components of interest is varied. Alternatively the SLICING method may be used for unique resolvation using only one sample (Engelsen and Bro, 2003, Manetti et al., 2004).

5. Conclusions and perspectives

This work has been concerned with various aspects of wheat quality for bread making. The process from grain to bread has been assessed on three major levels of different perspective. A basic study regarding the fundamental heterogeneity of particulate materials such as bulk wheat has been performed with respect to truly representative sampling for estimating the true average property or concentration of an unknown critical component of the lot, e.g. the protein content (Paper I). The fundamental heterogeneity or diversity of bulk wheat was utilised in a holistic perspective of the entire baking process from grain to the sensory perception of bread texture by the human brain on bulk wheat diversified by sorting according to an internal complex quality trait (Paper II and III). A zoom on water and fat compartmentalisation in bread dough lead to an investigation of the analytical properties of the current 2D-Lapace inversion technique and the development of a novel chemometric method for analysing diffusion-relaxation correlation spectra in nuclear magnetic resonance spectroscopy (Paper IV).

Paper I demonstrated the applicability of Pierre Gy's theory of sampling within bulk wheat and serves as a reference study for quantifications of the sampling errors involved in all attempts to estimate true average properties of composite materials not limited to protein content in wheat. The fundamental, the grouping and segregation and the analytical errors contributing to the global estimation error were quantified. The sampling errors were an order of magnitude larger than the analytical errors in composite samples of forty-two seeds and thus stress the importance of true replicates of the entire sampling process for assessing the variance of the global estimation error. The comparison of grab sampling with truly representative sampling by riffle splitting showed that grab sampling may lead to erroneous variance estimates – either too high or too low as well as biases on the analytical result of unknown size. The investigation also contributed to support the feasibility of utilising the inherent heterogeneity for improving bulk market value by sorting according to important properties such as the protein content.

Paper II and III showed that sorting grains post harvest with respect to an internal complex quality trait using single-kernel near-infrared (SKNIR) technology was very effective for diversification of organically grown wheat as compared to diversification based on different preceding catch crops tested in two growth years and two locations/cultivars. The SKNIR sorting improved the protein and wet gluten contents markedly. The Zeleny sedimentation volume, Farinograph water absorption and falling number were likewise improved. The bulk grain density of the best quality fraction was also increased compared to the starting material. The potential added value was apparently only limited by the constitutional heterogeneity of the bulk wheat.

In the prediction of a number of flour functionality parameters near-infrared reflectance spectroscopy on flour was found superior as compared to NIT on grain, NIT on flour, IR on flour and NMR relaxometry on the baking process and combinations thereof. However, the flour functionality parameters as predictors were themselves superior to all the spectroscopic techniques for predicting bread quality in terms of dimensions and texture. The α-amylase activity, Farinograph development time and softening and gluten index were all very important variables for the prediction of volume and texture parameters; hardness, chewiness and fracturability. Farinograph water absorption was important for prediction of bread mass and sensory perceived elasticity and dryness. It was also shown that the automatic mixing function in the baking test using an automatic home-bakery should be avoided since an unfavourable preference for weak doughs by the machine technology may mask effects of varying protein quantity and quality.

In Paper IV a novel method for analysing 2D diffusion-relaxation correlation data from a number of samples with varying properties of interest was proposed and tested on bread doughs. The method combines two well known methods; PARAFAC and 2D-Laplace inversion and solves a number challenges in this type of data, not limited to water and fat compartmentalisation and dynamics in dough. The PARAFAC resolves 2D diffusion-relaxation multi-exponentially decaying landscapes into unique quantifiable components which are subsequently translated to the T_2-D domain by Laplace inversion for interpretation. Furthermore the concatenated procedure simplifies the determination of defining appropriate regularisation of the labile Laplace inversion and filters artefacts of individual diffusion-relaxation re-

cordings which might else be interpreted as true components in the T_2-D domain.

The four papers and the present thesis thus examine a few but very important aspects of the complex process of bread making. The concepts of sampling, chemometric technology for sorting and prediction of end product properties as well as chemometric method developments however reach far beyond the applications in wheat, dough and bread presented herein.

Dedicated equipment for sampling in various materials will be developed and tested in order to ensure just estimations of critical properties and methods routinely used in test laboratories today will be considered bad laboratory practices in the future. A process which will be driven by customer demand for trustworthy analytical results and substantiated by heterogeneity characterisations and quantification of hidden errors generated in present procedures. The responsibility of the sampling process will move into the domain of the test laboratories and not rely on simplified instructions on how to prepare the sample prior to shipment and analysis. Although the importance of sampling processes have been known for decades especially in certain industries where the analytical results have been central for running the operation, this quest is only now beginning in areas such as environmental studies, biology and food science and industry.

Industrial and consumer demands for diversified quality product can at least in some areas be met by sorting on internal quality as shown for wheat as long as the fragments sorted have a reasonable macroscopic size for practical handing by the equipment. However, the theoretical as well as the practical size limit for handling individual fragments is on the molecular scale, known as nano-technology.

New methods for analysing chemical and physical phenomena are continuously being developed which will demand new user-friendly mathematical, statistical and chemometric methods for interpretation and quantification of the observations. The enormous amount of data being collected everywhere is generating a new frontier of data management and treatment to ensure manageable traceability and possibilities of extraction and analysis without endless conversions with the risk of trace destruction and confusions.

The integrated official initiative by the United States Food and Drug Administration for implementing the above mentioned process analytical technologies in the pharmaceutical industries is foreseen to have marked effects on research and developments in a wide range of industries as well as in a number of natural sciences.

6. References

Abe, H., Kusama, T., Kawano, S., Iwamoto, M. (1996): Non-destructive determination of protein content in a single kernel of wheat and soybean by near-infrared spectroscopy. In: Davies, A.M.C., Williams, P. (eds.): *Near-Infrared Spectroscopy: Future Waves*. NIR Publications, Chichester, UK, 457-461.

Allen, T., Khan, A.A. (1970): Critical evaluation of powder sampling procedures. *The Chemical Engineer*, 238, 108-112.

Armstrong, P.R. (2006): Rapid single-kernel NIR measurement of grain and oil-seed attributes. *Applied Engineering in Agriculture*, 22(5), 767-772.

Bakeev, K.A. (2005): *Process analytical technology*, Blackwell Publishing, Oxford, UK.

Barnes, R.J., Dhanoa, M.S., Lister, S.J. (1989): Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, 43(5), 772-777.

Belderok, B. (2000): Developments in bread-making processes. *Plant Foods for Human Nutrition*, 55, 1-14.

Belitz, H.-D., Grosch, W. (1999): *Food Chemistry. Second Edition*. Springer-Verlag, Berlin.

Bourne, M.C. (2002): *Food texture and viscosity: concept and measurement. Second Edition.* Academic Press, London.

Bourne, M.C., Comstock, S.H. (1981): Effect of degree of compression on texture parameters. *Journal of Texture Studies*, 12, 201-216.

Brás, L.P., Bernardino, S.A., Lopes, J.A., Menezes, J.C. (2005): Multiblock PLS as an approach to compare and combine NIR and MIR spectra of soybean flour. *Chemometrics and Intelligent Laboratory Systems*, 75, 91-99.

Bro, R. (1996): *Håndbog i multivariat kalibrering*. DSR Forlag, Frederiksberg, Denmark. In Danish.

Bro, R. (1997): PARAFAC. Tutorial and application. *Chemometrics and Intelligent Laboratory Systems*, 38(2), 149-171.

Bro, R., van den Berg, F., Thybo, A., Andersen, C.M., Jørgensen, B.M., Andersen, H. (2002): Multivariate data analysis as a tool in advanced quality monitoring in the food production chain. *Trends in Food Science & Technology*, 13, 235-244.

Brown, W.H. (1995): *Organic Chemistry*. International Edition. Saunders College Publishing. FL. USA.

Bushuk, W., Briggs, K.G., Shebeski, L.H. (1969): Protein quantity and quality as factors in the evaluation of bread wheats. *Canadian journal of plant science*, 49(2), 113-122.

Butler, J.P., Reeds, J.A, Dawson, S.V. (1981): Estimating solutions of 1st kind integral-equations with nonnegative constraints and optimal smoothing. *SIAM Journal of Numeric Analysis*, 18, 381-397.

Callaghan, P.T. (1991): *Principles of Nuclear Magnetic Resonance Microscopy*, Clarendon Press, Oxford.

Callaghan, P.T., Godefroy, S., Ryland, B.N., (2003): Use of the second dimension in PGSE NMR studies of porous media. *Magnetic Resonance Imaging*, 21, 243–248.

Carr, H.Y., Purcell, E.M. (1954): Effect of diffusion on free precession in nuclear magnetic resonance experiments. *Physical Reviews*, 94, 630-638.

Carroll, J.D., Chang, J. (1970): Analysis of individual differences in multidimensional scaling via N-way generalization and Eckart-Young decomposition, *Psychometrika*, 35, 283-319.

Choi, S.W., Lee, I.-B. (2005): Multiblock PLS-based localized process diagnosis. *Journal of Process Control*, 15, 295-306.

Coates, J.P. (2005): Infrared spectroscopy for process analytical applications. In: Bakeev, K.A. (ed.): *Process analytical technology*, Blackwell Publishing, Oxford, UK.

Coblentz, W.W. (1905): *Investigations of infrared spectra*, Carnegie Institution of Washington, Washington, D.C.

Dahm, D.J., Dahm, K.D. (2001): The physics of near-infrared scattering. In: Williams, P., Norris, K. (eds.) (2001): *Near-infrared technology in the agricultural and food industries, 2nd edition*, AACC Press.

Delwiche, S.R. (1993): Measurement of single-kernel wheat hardness using near-infrared transmittance. *Transactions of the ASAE*, 36(5), 1431-1437.

Delwiche, S.R. (1995): Single wheat kernel analysis by near-infrared transmittance: Protein-content. *Cereal Chemistry*, 72(1), 11-16.

Delwiche S.R. (1998): Protein content of single kernels of wheat by near-infrared reflectance spectroscopy. *Journal of Cereal Science*, 27(3), 241-254.

Delwiche, S.R:, Weaver, G. (1994): Bread quality of wheat-flour by near-infrared spectrophotometry: Feasibility of modelling. *Journal of Food Science*, 59(2), 410-415.

Delwiche, S.R., Graybosch, R.A., Peterson, C.J. (1998): Predicting protein composition, biochemical properties, and dough-handling properties of hard red winter wheat flour by near-infrared reflectance. *Cereal Chemistry*, 75(4), 412-416.

Delwiche, S.R., Hruschka, W.R. (2000): Protein content of bulk wheat from near-infrared reflectance of individual kernels. *Cereal Chemistry*, 77(1), 86-88.

Delwiche, S.R., Graybosch, R.A. (2003): Examination of spectral pretreatments for partial least-squares calibrations for chemical and physical properties of wheat. *Applied Spectroscopy*, 57(12), 1517-1527.

Delwiche, S.R., Pearson, T.C., Brabec, D.L. (2005): High-speed optical sorting of soft wheat for reduction of deoxynivalenol. *Plant Disease*, 89(11), 1214-1219.

Devaux, M.F., Bertrand, D., Martin, G. (1986): Discrimination of bread-baking quality of wheats according to their variety by near-infrared reflectance spectroscopy. *Cereal Chemistry*, 63(2), 151-154.

Dowell, F.E. (2000): Differentiating vitreous and nonvitreous durum wheat kernels by using near-infrared spectroscopy. *Cereal Chemistry*, 77(2), 155-158.

Dowell, F.E., Maghirang, E.B., Graybosch, R.A., Baenziger, P.S., Baltensperger, D.D., Hansen, L.E. (2006a): An automated near-infrared system for selecting individual kernels based on specific quality characteristics. *Cereal Chemistry*, 83(5), 537-543.

Dowell, F.E., Maghirang, E.B., Xie, F., Lookhart, G.L., Pierce, R.O., Seabourn, B.W., Bean, S.R., Wilson, J.D., Chung, O.K. (2006b): Predicting wheat quality characteristics and functionality using near-infrared spectroscopy. *Cereal Chemistry*, 83(5), 529-536.

Eckart, C., Young, G. (1936): The approximation of one matrix by another of lower rank. *Psychometrika*, 1, 211-218.

Elfverson, C., Andersson, A.A.M., Åman, P., Regnér, S. (1999): Chemical composition of barley cultivars fractioned by weighing, pneumatic classification, sieving, and sorting on a specific gravity table. *Cereal Chemistry*, 76(3), 434-438.

Engelsen, S.B., Jensen, M.K, Pedersen, H.T., Nørgaard, L., Munck, L. (2001): NMR-baking and multivariate prediction of instrumental texture parameters in bread. *Journal of Cereal Science*, 33, 59-69.

Engelsen, S.B., Bro, R. (2003): Powerslicing. *Journal of Magnetic Resonance*, 163(1), 192-197.

Esbensen, K.H. (2000): *Multivariate data analysis - in practice. 4th edition*. Camo, Oslo, Norway.

Esbensen, K., Geladi, P. (1990): The start and early history of chemometrics: Selected interviews. Part 2. *Journal of Chemometrics*, 4, 355-412.

Esbensen, K.H., Friis-Petersen, H.H., Petersen, L., Holm-Nielsen, J.B., Mortensen, P.P. (2007): Representative process sampling – in practice: Variographic analysis and estimation of total sampling errors (TSE), *Chemometrics and Intelligent Laboratory Systems*, 88(1), 41-59.

Fisher, R.A. (1950): *Statistical methods for research workers, Eleventh edition*. Oliver and Boyd, Edinburgh.

Fontaine, J., Schirmer, B., Hörr, J. (2002): Near-Infrared Reflectance Spectroscopy (NIRS) enables the fast and accurate prediction of essential amino acid contents. 2. Results for wheat, barley, corn, triticale, wheat bran/middlings, rice bran, and sorghum. *Journal of Agricultural and Food Chemistry*, 50, 3902-3911.

Friedman, H.H, Whitney, J.E, Szczesniak, A.S. (1963): The texturometer – a new instrument for objective texture measurement. *Journal of food science*, 28, 390-396.

Gabrielsson, J., Trygg, J. (2006): Recent developments in multivariate calibration. *Critical Reviews in Analytical Chemistry*, 36(3-4), 243-255.

Geladi, P., MacDougall, D., Martens, H. (1985): Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Applied Spectroscopy*, 39(3), 491-500.

Geladi, P., Kowalski, B.R. (1986): Partial least-squares regression: A tutorial. *Analytica Chimica Acta*, 185, 1-17.

Geladi, P., Esbensen, K. (1990): The start and early history of chemometrics: Selected interviews. Part 1. *Journal of Chemometrics*, 4, 337-354.

Gerlach, R.W., Dobb, D.E., Raab, G.A., Nocerino, J.M. (2002): Gy sampling theory in environmental studies. 1. Assessing soil splitting protocols. *Journal of Chemometrics*, 16, 321-328.

Gerlach, R.W., Nocerino, J.M., Ramsey, C.A., Venner, B.C. (2003): Gy sampling theory in environmental studies. 2. Subsampling error estimates. *Analytica Chimica Acta*, 490, 159-168.

Godefroy, S., Callaghan, P.T. (2003): 2D relaxation/diffusion correlations in porous media. *Magnetic Resonance Imaging*, 21, 381–383.

Golub, G.H., Reinsch. C. (1970): Singular value decomposition and least-squares solutions. *Numerische Mathematik*, 14(2), 403-420.

Grausgruber, H., Kreuzmayr, A.E., Ruckenbauer, P. (2001): Evaluation of the breadmaking quality of Austrian-grown wheats using an automatic home-bakery. *Cereal Research Communications*, 29(3-4), 421-428.

Graybosch, R., Peterson, C.J., Moore, K.J., Stearns, M., Grant, D.L. (1993): Comparative effects of wheat-flour protein, lipid, and pentosan composition in relation to baking and milling quality. *Cereal Chemistry*, 70(1), 95-101.

Greffeuille, V., Abecassis, J., Rousset, M., Oury, F.X., Faye, A., Bar L'Helgouac'h, C., Lullien-Pellerin, V. (2006): Grain characterization and milling behavior of near-isogenic line differing by hardness. *Theoretical and Applied Genetics*, 114, 1-12.

Guenard, R., Thurau, G. (2005): Implementation of process analytical technologies. In: Bakeev, K.A. (ed): *Process analytical technology*, Blackwell Publishing, Oxford, UK.

Gy, P. (1986): The analytical and economic importance of correctness in sampling. *Analytica Chimica Acta*, 190, 13-23.

Gy, P. (1995a): Introduction to the theory of sampling. I. Heterogeneity of a population of uncorrelated units. *Trends in Analytical Chemistry*, 14(2), 67-76.

Gy, P. (1995b): The future of analysis: To check the sampling errors is the only way to improve the overall analytical reliability. *Analusis*, 23(10), 497-500.

Gy, P. (1998): *Sampling for analytical purposes*, John Wiley and Sons.

Gy, P. (2004a): Sampling of discrete materials – a new introduction to the theory of sampling – I. Qualitative approach. *Chemometrics and Intelligent Laboratory Systems*, 74, 7-24.

Gy, P. (2004b): Sampling of discrete materials – II. Quantitative approach. *Chemometrics and Intelligent Laboratory Systems*, 74, 25-38.

Gy, P. (2004c): Sampling of discrete materials – III. Qualitative approach – sampling of one-dimensional objects. *Chemometrics and Intelligent Laboratory Systems*, 74, 39-47.

Gy, P. (2004d): Part IV: 50 years of sampling theory – a personal history. *Chemometrics and Intelligent Laboratory Systems*, 74, 49-60.

Hansen, B., Hansen, Å. (1992): Test baking of bread by household baking machine. *Food Science and Technology*, 25, 585-587.

Hansen, B., Hansen, Å. (1993): Erratum. Test baking of bread by household baking machine. *Food Science and Technology*, 26, 181.

Hardacre, A.K, (2006): Personal Communication. Crop & Food Research, Palmerston North, New Zealand.

Harshman, R.A. (1970): Foundation of the PARAFAC procedure: Model and conditions for an 'explanatory' multi-mode factor analysis. *UCLA Working Papers in Phonetics*, 16, 1-84.

Henry, W.F., Katz, M.H. (1969): New dimensions relating to the textural quality of semi-solid foods and ingredient systems. *Food Technology*, 23, 822-825.

Henry, W.F., Katz, M.H., Pilgrim, F.J., May, A.J. (1971): Texture of semi-solid foods: sensory and physical correlates. *Journal of food science*, 36, 155-161.

Herschel, W. (1800): Investigations of the powers of the prismatic colours to heat and illuminate objects; with remarks, that prove the different refrangibility of radient heat. To which is added, an inquiry into the method of viewing the sun advantageously, with telescopes of large apertures and high magnifying powers. *Philosophical Transactions of the Royal Society of London*, 90, 255-283.

Heydorn, K., Hansen, E.H. (2005): Does gravitational segregation of ions really exist? *Chemometrics and Intelligent Laboratory Systems*, 79(1-2), 129.

Holm-Nielsen, J.B., Dahl, C.K., Esbensen, K.H. (2006): Representative sampling for process analytical characterization of heterogeneous bioslurry systems – a reference study of sampling issues in PAT. *Chemometrics and Intelligent Laboratory Systems*, 83, 114-126.

Hotelling, H. (1933): Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, 417-441, 498-520.

Hubbard, P.L., Watkinson, P.J., Creamer, L.K., Gottwald, A., Callaghan, P.T. (2005): Two-dimensional Laplace inversion NMR technique applied to the molecular properties of water in dry-salted mozzarella-type cheeses with various salt concentrations. In: Engelsen, S.B., Belton, P.S., Jakobsen, H.J. (eds.): *Magnetic resonance in food science. The multivariate challenge.* The proceedings of the 7th international conference on applications of magnetic resonance in food science held in Copenhagen on the 13–15th September 2004, The Royal Society of Chemistry, Cambridge, UK, pp. 225–232.

Huebner, F.R., Gaines, C.S. (1992): Relation between wheat kernel hardness, environment, and gliadin composition. *Cereal Chemistry*, 69(2), 148-151.

Hürlimann, M.D., Venkataramanan, L., Flaum, C. (2002): The diffusion-spin relaxation time distribution function as an experimental probe to characterize fluid mixtures in porous media, *Journal of Chemical Physics*, 117, 10223–10232.

Ingamells, O., Switzer, P. (1973): A proposed sampling constant for use in geochemical analysis. *Talanta*, 20, 547-568.

Istratov, A.A, Vyvenko, O.F. (1999): Exponential analysis of physical phenomena. *Review of Scientific Instruments*, 70(2), 1233-1257.

Jestel, N.L. (2005): Process Raman spectroscopy. In: Bakeev, K.A. (ed.): *Process analytical technology*, Blackwell Publishing, Oxford, UK.

Kaye, B.H. (1967): Sampling: Cinderella of the analyst's art. In: *Rock Products*. Illinois Institute of Technology.

Kourti, T. (2006): Process analytical technology beyond real-time analyzers: The role of multivariate analysis. *Critical reviews in Analytical Chemistry*, 36, 257-278.

Kowalski, B.R. (1975): Chemometrics: Views and propositions. *Journal of Chemical Information and Computer Science*, 15(4), 201-203.

Kowalski, B., Gerlach, R., Wold, H. (1982): Chemical systems under indirect observation. In: Jöreskog, K.G., Wold, H. (eds.): *Systems under indirect observation II*, North-Holland, Amsterdam, 191-209.

Kroonenberg, P.M. (1997): *Introduction to biplots for G×E tables, Version 3*. Research reports #51, Centre for Statistics, The University of Queensland, Brisbane, Australia.

Kuhn, T.S. (1970): *The structure of scientific revolutions. 2nd edition*. University of Chicago Press.

Lamé, F., Honders, T., Derksen, G., Gadella, M. (2005): Validated sampling strategy for assessing contaminants in stockpiles. *Environmental Pollution*, 134, 5-11.

Larsen, J.U. (1999): *Fremtidens brød af fortidens korn*. Forlaget Olivia, Denmark. In Danish.

Leardi, R. (2001): Genetic algorithms in chemometrics and chemistry: a review. *Journal of Chemometrics*, 15, 559-569.

Löfqvist, B., Nielsen, J.P. (2003): A method for sorting objects comprising organic material. *Patent* WO 03/004179 A1.

Löfqvist, B., Nielsen, J.P. (2004): Method and device for sorting objects. *Patent* WO 2004/060585 A1.

Lwin, T., Flann, R.C.A., Short, G.M., Guthrie, W. (1998): Design and analysis of size-mass reduction experiments for sampling particulate materials. *International Journal of Mineral Processing*, 54, 59-80.

Maghirang, E.B., Dowell, F.E. (2003): Hardness measurement of bulk wheat by single-kernel visible and near-infrared reflectance spectroscopy. *Cereal Chemistry*, 80(3), 316-322.

Manetti, C., Castro, C., Zbilut, J.P. (2004): Application of trilinear SLICING to analyse a single relaxation curve. *Journal of Magnetic Resonance*, 168(2), 273-277.

Martens, H., Wedøe, S., Martens, M. (no year): *Guide to sensory data analysis in the Unscrambler*. Sensory Science Group, Dept. of Food Science and Technology, The Royal Veterinary and Agricultural University, Copenhagen, Denmark.

Martens, H., Stark, E. (1991): Extended multivariate signal correction and spectral interference subtraction: new preprocessing methods for near infrared spectroscopy. *Journal of Pharmaceutical and Biomedical Analysis*, 9(8), 625-635.

Martens, H., Martens, M. (2000): Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Quality and Preference*, 11, 5-16.

Martens, H., Nielsen, J.P., Engelsen, S.B. (2003): Light scattering and light absorbance separated be extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures, *Analytical Chemistry*, 75(3), 394-404.

Martens, M., Bredie, W.L.P., Martens, H. (2000): Sensory profiling data studied by partial least squares regression, *Food Quality and Preference*, 11, 147-149.

McNaught, A.D., Wilkinson, A. (1997): *IUPAC Compendium of Chemical Terminology. The Gold Book, 2nd Edition*. Blackwell Science. [http://goldbook.iupac.org/index.html], June 2007.

Meiboom, S., Gill, D. (1958): Modified spin-echo method for measuring nuclear relaxation times. *The Review of Scientific Instruments*, 29, 688-691.

Meilgaard, M., Civille, G.V., Carr, B.T. (1991): Introduction to sensory techniques. In: *Sensory Evaluation Techniques*, CRC Press, Boca Raton, FL, USA, 1-22.

Micklander, E., Peshlov, B., Purslow, P.P., Engelsen, S.B. (2002): NMR-cooking: monitoring the changes in meat during cooking by low-field ^{1}H-NMR. *Trends in Food Science & Technology*, 13, 341-346.

Micklander, E., Thygesen, L.G., Pedersen, H.T., van den Berg, F., Bro, R., Rutledge, D.N., Engelsen, S.B. (2003): Multivariate analysis of time domain NMR signals in relation to food quality. In: Belton, P.S., Webb, G.A., Gil, A.M., Rutledge D.N (eds.): *Magnetic Resonance in Food Science: Latest developments*. Royal Society of Chemistry, UK, 239-254.

Miralbés, C. (2003): Prediction chemical composition and Alveograph parameters on wheat by near-infrared transmittance spectroscopy. *Journal of Agricultural and Food Chemistry*, 51, 6335-6339.

Miralbés, C. (2004): Quality control in the milling industry using near infrared transmittance spectroscopy. *Food Chemistry*, 88, 621-628.

Munck, L. (2005): *The revolutionary aspect of chemometric technology*. Narayana Press, Gylling, Denmark.

Munck, L. (2008): Breeding for quality traits in cereals – a revised outlook on old and new tools for integrated breeding. In: Carena M.J. (Ed.): *Cereals*. Springer Verlag, NY, in print.

Nielsen, J.P. (2002): *Fast quality assessment of barley and wheat: Chemometric exploration of instrumental data with single seed applications*. Ph.D. thesis. Food Technology, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Frederiksberg, Denmark.

Nielsen, J.P., Pedersen, D.K., Munck, L. (2003): Development of nondestructive screening methods for single kernel characterization of wheat. *Cereal Chemistry*, 80(3), 274-280.

Nørgaard, L., Saudland, A., Wagner, J., Nielsen, J.P., Munck, L., Engelsen, S.B. (2000): Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. *Applied Spectroscopy*, 54(3), 413-419.

Osborne, B.G. (1984): Investigations into the use of near-infrared reflectance spectroscopy for the quality assessment of wheat with respect to its potential for bread baking. *Journal of the Science of Food and Agriculture*, 35, 106-110.

Osborne, B.G. (2006): Applications of near infrared spectroscopy in quality screening of early-generation material in cereal breeding programmes. *Journal of Near Infrared Spectroscopy*, 14, 93-101.

Osborne, B.G., Douglas, S. (1981): Measurement of the degree of starch damage in flour by near-infrared reflectance analysis. *Journal of the Science of Food and Agriculture*, 32(4), 328-332.

Osborne, B.G., Douglas, S., Fearn, T. (1982): The application of near-infrared reflectance analysis to rapid flour testing. *Journal of Food Technology*, 17(3), 355-363.

Osborne, B.G., Fearn, T., Hindle, P.H. (1993): *Practical NIR spectroscopy with applications in food and beverage analysis*. Longman Scientific and Technical, Harlow.

Pasikatan, M.C., Dowell, F.E. (2004): High-speed NIR segregation of high- and low-protein single wheat seeds. *Cereal Chemistry*, 81(1), 145-150.

Pearson, K. (1901): On lines and planes of closest fit to systems of points in space, *Philosophical Magazine*, 2, 559-572.

Pedersen, D.K. (2002): *Spectroscopic and chemometric exploration of food quality*. Ph.d. thesis. Food Technology, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Frederiksberg, Denmark.

Pedersen, D.K., Engelsen, S.B. (2001): Monitoring industrial food processes using spectroscopy and chemometrics. *New Food*, 2, 9-13.

Pedersen, D.K., Martens, H., Nielsen, J.P., Engelsen, S.B. (2002a): Near-infrared absorption and scattering separated by extended inverted signal correction (EISC): Analysis of near-infrared transmittance spectra of single wheat seeds. *Applied Spectroscopy*, 56(9), 1206-1214.

Pedersen, H.T., Munck, L., Engelsen, S.B. (2000): Low field H-1 nuclear magnetic resonance and chemometrics combined for simultaneous determination of water, oil, and protein contents in oilseeds. *Journal of the American Oil Chemists Society*, 77(10), 1069-1076.

Pedersen, H.T., Bro, R., Engelsen, S.B. (2002b): Towards rapid and unique curve resolution of low-field NMR relaxation data: Trilinear SLICING versus two-dimensional curve fitting. *Journal of Magnetic Resonance*, 157(1), 141-155.

Peltonen, J., Salovaara, H. (1991): Experiences of an automatic small-scale home bakery in test bakings of 6 spring wheat-varieties. *Journal of Agricultural Science in Finland*, 63(2), 131-135.

Petersen, L., Dahl, C.K., Esbensen, K.H. (2004): Representative mass reduction in sampling – a critical survey of techniques and hardware. *Chemometrics and Intelligent Laboratory Systems*, 74, 95-114.

Petersen, L., Esbensen, K.H. (2005): Representative process sampling for reliable data analysis – a tutorial. *Journal of Chemometrics*, 19, 625-647.

Petersen, L., Minkkinen, P., Esbensen, K.H. (2005): Representative sampling for reliable data analysiss: Theory of Sampling. *Chemometrics and Intelligent Laboratory Systems*, 77, 261-277.

Pinckney, A.J., Greenaway, W.T., Zeleny, L. (1957): Further developments in the sedimentation test for wheat quality, *Cereal Chemistry*, 34(1), 16-25.

Pitard, F.F. (1993): *Pierre Gy's sampling theory and sampling practice. Heterogeneity, sampling correctness, and statistical process control*, CRC Press.

Qiao, Y., Galvosas, P., Callaghan, P.T. (2005): Diffusion correlation NMR spectroscopy study of anisotropic diffusion of water in plant tissues. *Biophysical Journal*, 89(4), 2899–2905.

Qin, J.S, Valle, S., Piovoso, M.J. (2001): On unifying multiblock analysis with application to decentralized process monitoring. *Journal of Chemometrics*, 15, 715-742.

Rittiron, R., Saranwong, S., Kawano, S. (2004): Useful tips for constructing a near infrared-based quality sorting system for single brown-rice kernels. *Journal of Near Infrared Spectroscopy*, 12(2), 133-139.

Rubenthaler, G.L., Bruinsma, B.L. (1978): Lysine estimation in cereals by near-infrared reflectance. *Crop Science*, 18(6), 1039-1042.

Samp, E.J., Sedin, D., Foster, A. (2003): Enhanced NIR calibration for wort fermentability using orthogonal signal correction. *Journal of the Institute of Brewing*, 109(1), 16-26.

Savitsky, A., Golay, M.J.E. (1964): Smoothing and differentiation of data by simplified least squares procedures, *Analytical Chemistry*, 36(8), 1627-1639.

Siesler, H.W., Ozaki,Y., Kawata, S., Heise, H.M. (eds.) (2002): *Near-infrared spectroscopy, principles, instruments, applications*. WILEY-VCH Verlag GmbH, Weinheim, Germany.

Skovgaard, I.M., (1996): *Statistisk forsøgsplanlægning*. DSR Forlag, Copenhagen. In Danish.

Smith, P.L. (2004): Audit and assessment of sampling systems. *Chemometrics and Intelligent Laboratory Systems*, 74, 225-230.

Song, Y.-Q., Venkataramanan, L., Hürlimann, M.D., Flaum, M., Frulla, P., Straley, C. (2002): T1–T2 correlation spectra obtained using fast twodimensional laplace inversion. *Journal of Magnetic Resonance*, 154, 261–268.

Svensson, O., Kourti, T., MacGregor, J.F. (2002): An investigation of orthogonal signal correction algorithms and their characteristics. *Journal of Chemometrics*, 16(4), 176-188.

Szczesniak, A.S. (1963a): Classification of textural characteristics. *Journal of food science*, 28(4), 385-389.

Szczesniak, A.S. (1963b): Objective measurements of food texture. *Journal of food science*, 28(4), 410-420.

Szczesniak, A.S. (1968): Correlations between objective and sensory measurements. *Food Technology*, 22, 981-985.

Szczesniak, A.S., Brandt, A.M., Friedman, H.H. (1963): Development of standard rating scales for mechanical parameters of texture and correlation between the objective and sensory methods of texture evaluation. *Journal of Food Science*, 28, 397-403.

Trygg, J. (2002): O2-PLS for qualitative and quantitative analysis in multivariate calibration. *Journal of Chemometrics*, 16(6), 283-293.

Trygg, J., Wold, S. (2002): Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics*, 16(3), 119-128.

Trygg, J., Wold, S. (2003): O2-PLS, a two-block (X-Y) latent variable regression (LVR) method with an integral OSC filter. *Journal of Chemometrics*, 17(1), 53-64.

USFDA (2004a): *Guidance for industry. PAT – A framework for innovative pharmaceutical development, manufacturing and quality assurance*. US Department of Health and Human Services, Center for Drug Evaluation and Research (CDER), Rockville, MD 20857, United States.
[http://www.fda.gov/cder/guidance/6419fnl.pdf], June 2007.

USFDA (2004b): *Pharmaceutical CGMPs for the 21st century – A risk-based approach*. Department of Health and Human Services, Center for Drug Evaluation and Research (CDER), Rockville, MD 20857, United States.
[http://www.fda.gov/cder/gmp/gmp2004/CGMP report final04.pdf], June 2007.

USFDA (2005): *Process analytical technology (PAT) initiative*. [http://www.fda.gov/cder/ops/pat.htm], June 2007

Venables, H.J., Wells, J.I. (2002): Powder sampling. *Drug Development and Industrial Pharmacy*, 28(2), 107-117.

Veraverbeke, W.S., Delcour, J.A. (2002): Wheat protein composition and properties of wheat glutenin in relation to breadmaking functionality. *Critical Reviews in Food Science and Nutrition*, 42(3), 179-208.

Verron, T., Sabatier, R., Joffre, R. (2004): Some theoretical properties of the O-PLS method. *Journal of Chemometrics*, 18(2), 62-68.

Vold, R.L., Waugh, J.S., Klein, M.P., Phelps, D.E. (1968): Measurement of spin relaxation in complex systems. *Journal of Chemical Physics*, 48(8), 3831-3832.

Wesley, I.J., Larroque, O., Osborne, B.G., Azudin, N., Allen, H., Skerritt, J.H. (2001): Measurement of gliadin and glutenin content of flour by NIR spectroscopy. *Journal of Cereal Science*, 34, 125-133.

Westerhuis, J.A., Coenegracht, P.M.J. (1997): Multivariate modelling of the pharmaceutical two-step process of wet granulation and tableting with multiblock partial least squares. *Journal of Chemometrics*, 11, 379-392.

Westerhuis, J.A, Kourti, T., Macgregor, J.F. (1998): Analysis of multiblock and hierarchical PCA and PLS models. *Journal of Chemometrics*, 12, 301-321.

Westerhuis, J.A, Smilde, A.K. (2001): Deflation in multiblock PLS. *Journal of Chemometrics*, 15, 485-493.

Whitaker, T.B., Halger, W.M., Giesbrecht, F.G., Johansson, A.S. (2000): Sampling, sample preparation, and analytical variability associated with testing wheat for deoxynivanol. *Journal of AOAC International*, 83(5), 1285-1292.

Williams, H., Flemming, I. (1995): *Spectroscopic methods in organic chemistry, fifth edition*. McGraw-Hill Publishing Company, England

Williams, P., Norris, K. (eds.) (2001): *Near-Infrared technology in the agricultural and food industries, 2nd edition*. American Association of Cereal Chemists, Inc. Minnesota, USA.

Wilson A.D. (1964): The sampling of rock powder for chemical analysis. *The Analyst*, 89, 18-30.

Wold, H. (1966): Nonlinear estimation by iterative least square procedures. In: David, F.N. (ed.): *Research papers in statistics, Festschrift for J. Neyman*, Wiley, New York, 411-444.

Wold, H. (1975): Soft modelling by latent variables: The non-linear iterative partial least squares (NIPALS) approach. In: Gani, J. (ed.): *Perspectives in probability and statistics*. Applied Probability Trust, Sheffield, England.

Wold, H. (1982): Soft modelling, the basic design and some extensions. In: Jöreskog, K.G., Wold, H. (eds.): *Systems under indirect observation II*, North-Holland, Amsterdam, 1-54.

Wold, S. (1991): Chemometrics, why, what and where to next? *Journal of Pharmaceutical and Biomedical Analysis*, 9(8), 589-596.

Wold, S. (1995): Chemometrics; what do we mean with it, and what do we want from it? *Chemometrics and Intelligent Laboratory Systems*, 30, 109-115.

Wold, S., Ruhe, A., Wold, H., Dunn, W.J. (1984): The collinearity problem in linear-regression - the partial least-squares (PLS) approach to generalized inverses. *Siam Journal on Scientific and Statistical Computing*, 5(3), 735-743.

Wold, S., Esbensen, K., Geladi, P. (1987): Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2, 37-52.

Wold, S., Antti. H., Lindgren, F., Öhman, J. (1998): Orthogonal signal correction of near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, 44(1-2), 175-185.

Workman, J., Koch, M., Veltkamp, D. (2005): Process analytical chemistry. *Analytical Chemistry*, 77(12), 3789-3806.

Yoon, B.S., Boroson, B.W., Lyford, C.P. (2002): Value of increasing kernel uniformity. *Journal of Agricultural and Resource Economics*, 27(2), 481-494.

Youden, W.J. (1951): *Statistical methods for chemists*. John Wiley and Sons, Inc, New York.

Zwingelberg, H., Brümmer, J.-M. (1990): Backautomaten in Mühlenlaboratorien. *Getreide, Mehl und Brot*, 44, 142-147.

Appendix A – Correction of sensory data

Due to the instability of the sensors in 'the sensory instruments' certain pre-processing evaluations and steps were necessary in the sensory data. The emphasis is on analysing the relative differences between samples rather than the absolute ratings given by the assessors and to remove attributes and assessors which do not contribute to the discrimination of the samples. In the following the pre-processing of the sensory data presented in Paper II and III is outlined in detail according to Martens et al. (2000) and Martens et al. (no year) guidance.

Histograms of ten sensory attributes determined by ten assessors in duplicates of the thirty-four (two more than presented in Paper II and III) bread samples are presented in Fig A1 and A2. Although data appear evenly distributed they do not show to what extent the assessors used the scales equally, nor whether they drift from replicate 1 to replicate 2. The aftertaste attribute (Fig. A1J) appears skewed in the distribution and might be working better after logarithmic transformation (Fig. A2). Although the log-corrected aftertaste had a better distribution, the original data were used in the subsequent data analysis. In Fig. A1 and the following four, attributes; yellowness, fracturability, hardness and chewiness were named colour, porosity, firmness and compactness respectively. They were renamed at a later stage, for better conceptual interpretation.

The **X** and **Y** loading of the level correction APLSR (Chap. 2.3) is presented in Fig. A3A. **Y** loadings in Fig. A3A show that highest level differences between judges and replicates are found in adhesiveness, chewiness (compactness), cereal aroma, wheat taste and fracturability (porosity). These differences were removed by utilising the **Y** residual after maximum PLS components (eleven) for further analysis (below). Note that only 9% of **X** can be explained by each component since all assessor-discriminating variables are orthogonal. However the APLSR-model itself is not very interesting, as it is only used as a method to remove level effects.

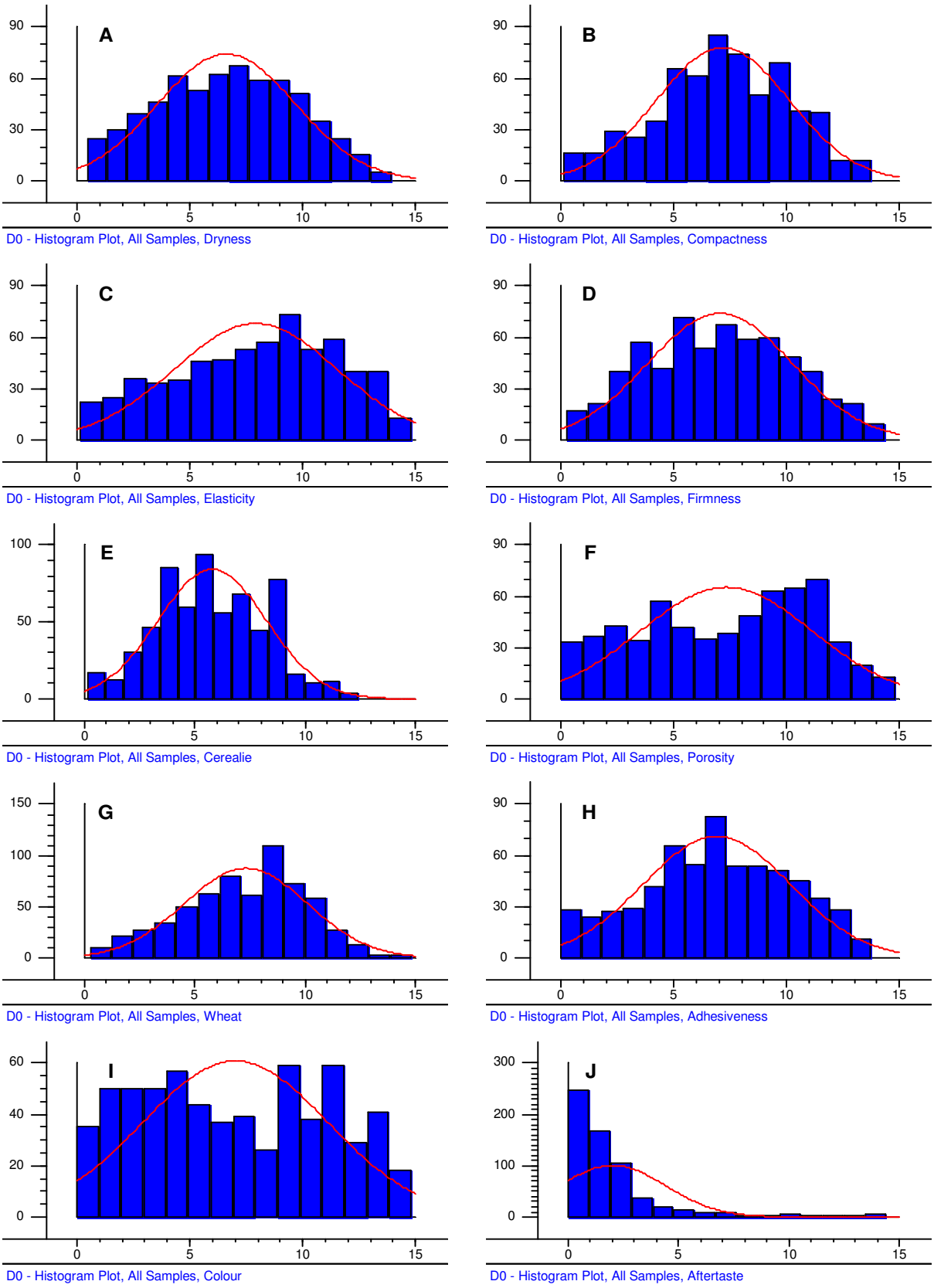


Figure A1: Histogram plots of 10 sensory attributes used in evaluation of bread samples showing that the entire sensory scale (0 – 15) has been utilised evenly across assessors and samples. In J, Aftertaste the distribution is clearly skewed and a logarithmic transformation might be appropriate (see Fig. A2)

In order to analyse assessor and attribute performances for discriminating individual bread samples an APLSR was calculated with the thirty-four different breads as **X** 0/1 design variables and the level corrected sensory data as **Y**. The model loadings can be seen in Fig. A3B.

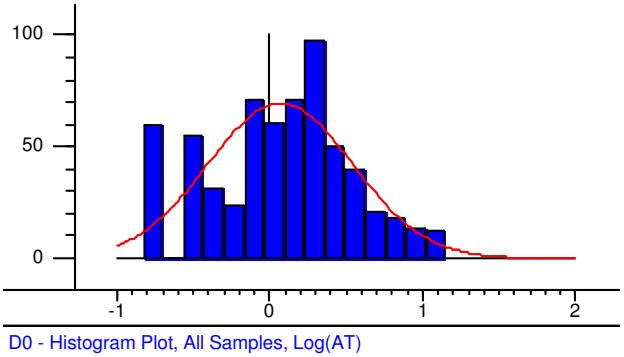


Figure A2: Log transformation of Aftertaste data ensures an even distribution of the attribute.

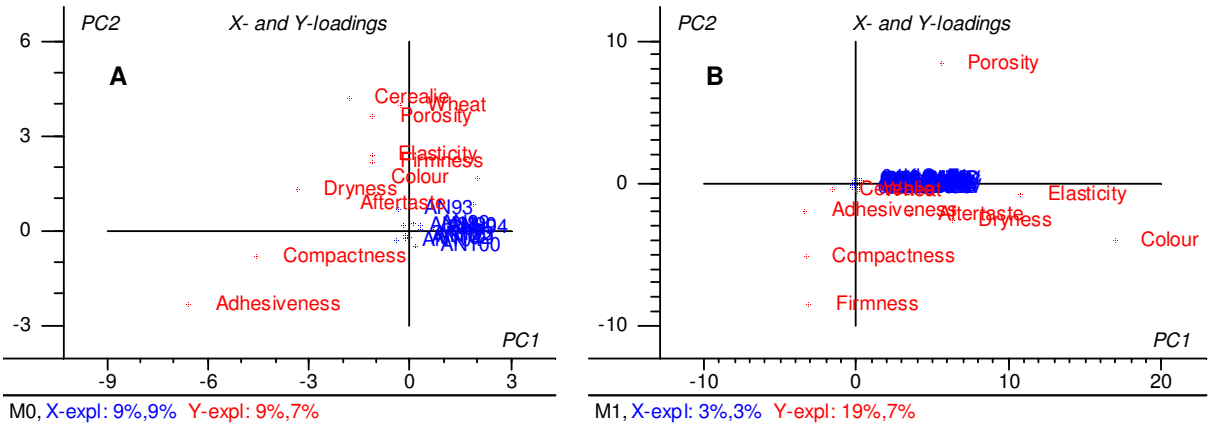


Figure A3: X- and Y-loadings plots of **A**: APLSR (level correction) of Assessors and replicates as **X** and sensory attributes as **Y**, **B**: APLSR of 34 products as **X** and level corrected sensory attributes as **Y**.

The products were best differentiated by yellowness (colour), elasticity, dryness and fracturability (porosity) on the first component, while hardness (firmness), fracturability (porosity) and chewiness (compactness) span the second component. Wheat and cereal tastes as well as adhesiveness and aftertaste do not contribute much to the differentiation of the samples.

In order to assess assessor and replication performances for discrimination of 34 bread samples the signal to noise ratio (S/N) is explored by comparing variation left after the optimal number of PCs (here 2) compared to average variations in the original data (Fig A4A). The original average variation in

signal levels along the x-axis indicates a variation in the use of the scales among the assessors, while the y-axis is the variation left after optimal number of PCs, i.e. the noise. Hence the S/N-ratios can be assessed graphically by the individual assessor's relative distances to the line indicating an S/N-ratio of 1. Assessor number 94 is generally using a much wider range of the sensory scale (highest) than assessor number 101 (lowest). However their S/N levels are in the same order of magnitude. Assessor number 93 indicated by an arrow generally have difficulties in discriminating samples.

The residuals for the individual sensory attributes are assessed similarly by averaging over sample and replicate variation at zero PCs and two PCs equal to no modelling and optimal model. In Fig. A4B the attributes yellowness (colour), elasticity and fracturability (porosity) are strongly contributing to the discrimination of individual samples. Dryness and hardness (firmness) also contribute, while cereal aroma and wheat taste are worthless for this investigation. The remaining attributes; chewiness (compactness), aftertaste and adhesiveness are less strong.

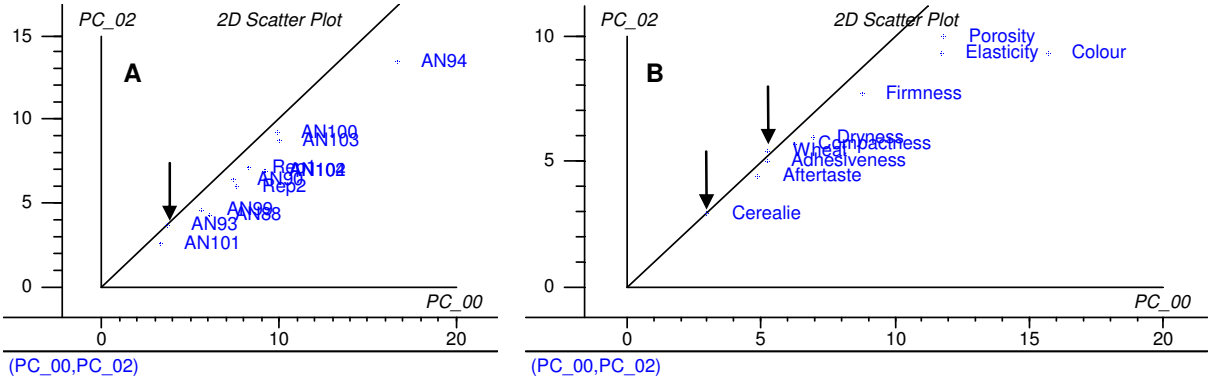


Figure A4: S/N-plots for assessors and replicates (A) and sensory attributes (B). PC_00 corresponds to average signal variation before APLSR modelling and PC_02 corresponds to the noise level (variation) after optimal number of PCs. Assessor number 93 and attributes cereal aroma and wheat tastes marked with arrows do not contribute to the discrimination of individual samples.

A detailed investigation of S/N plots for individual attributes (Fig. A5) showed that assessor number 93 indeed had low S/N ratio in all but yellowness (colour) attribute. Assessors 100 and 101 were having difficulties in five and six attributes respectively however their overall S/N ratios of 1.1 and 1.3 respectively in Fig. A4A and Fig A5F indicate that they are doing an overall reasonable job. The displays of five individual attributes in Fig A5A-E ex-

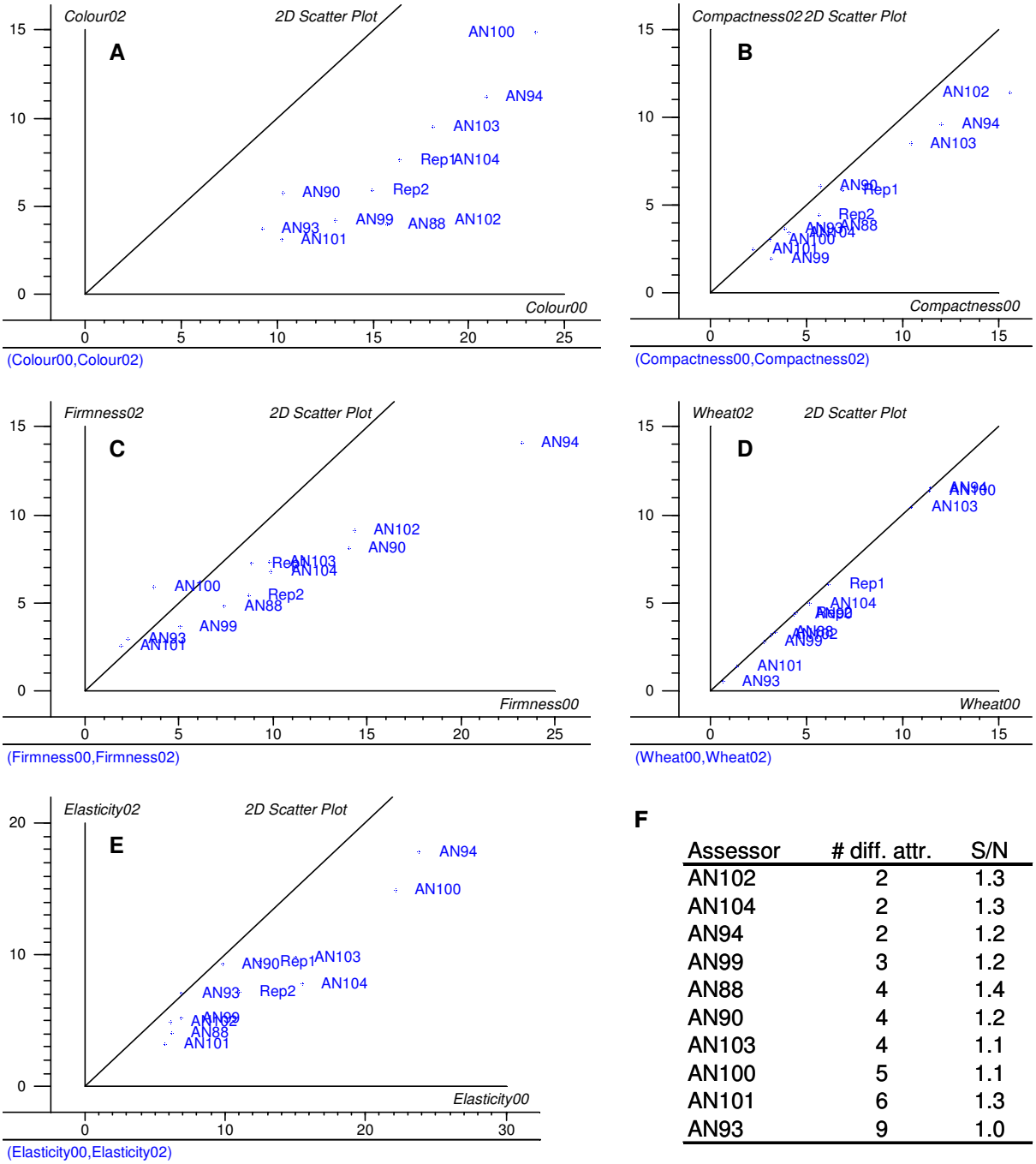
emplify the variation in the difficulties of determining certain attributes. While all assessors could differentiate well on yellowness (colour) (Fig. A5A), none could differentiate on wheat taste (Fig. A5D). In Fig A5B, C and E, the attributes chewiness (compactness), hardness (firmness) and elasticity were determined well by some assessors and bad by others. In summery the table included in Fig. A5F indicates that even though some assessors had difficulties in many attributes, their overall S/N ratio ensures their place in the panel.

The plots presented were utilised to make an informed selection of data for further analysis by expelling assessor number 93 data and attributes cereal aroma and wheat taste. Fig. A4B also supports the choice of not standardising the sensory attributes for the present analysis. Those attributes with high S/N ratios were also those with highest intensities. Weighting those down and those with low S/N ratios up would only complicate interpretation and on the same time affect the robustness negatively.

Discussion – why this might not have been the best approach
The procedure followed above evaluated assessors and attributes with respect to differentiating individual samples from the rest. Correlating thirty-four orthogonal directions to the ten sensory attributes with APLSR may not necessarily be the most appropriate strategy since a lot of the bread samples are actually expected to be quite similar, e.g. baked from lots originating from the same year and location/cultivar. The unfortunate choice is reflected in the low model complexity (two components) and low explained **X** and **Y** variations. Only one of 34 directions in **X** can be explained at a time, i.e. one sample as opposed to all the others, i.e. 3%. Thus only the most extreme samples are explaining the variations in **Y**, 19% on the first component and 7% on the second (Fig. A3B). Another approach could have been choosing a discriminator **X** matrix based on differences of special interest such as year, location/cultivar or fractionation differences. This would focus the assessor and attribute evaluations on their ability to discriminate special properties of interest.

Although the evaluation of assessors and attributes might not have been optimal, it does not affect the overall analysis presented in Paper II and III in a detrimental way. The sensory analyses of bread products using eight attributes were averaged over nine assessors and the two replicates. However

more attention to how data are corrected might point to new interesting features in the data not jet explored.



F

Assessor	# diff. attr.	S/N
AN102	2	1.3
AN104	2	1.3
AN94	2	1.2
AN99	3	1.2
AN88	4	1.4
AN90	4	1.2
AN103	4	1.1
AN100	5	1.1
AN101	6	1.3
AN93	9	1.0

Figure A5: Assessor S/N ratios for attributes A: Yellowness (colour), B: Chewiness (compactness), C: Hardness (firmness), D: Wheat taste and E: Elasticity. The table in F rank individual assessors performances by number of attributes determined with difficulties and overall S/N ratio (Fig. A4A).

Appendix B – The NIPALS algorithm with a PCA example

Although originally calculated differently by Pearson (1901), Hotelling (1933) and successors, the workhorse in actually calculating the principal components is the nonlinear iterative partial least squares (NIPALS) algorithm[11] developed by Herman Wold and co-workers in the 1960's and 70's (Wold, 1966, 1975).

The NIPALS algorithm:
1. Centre $\mathbf{X}$ across samples, i.e. subtract the mean of all variables in $\mathbf{X}$ from each row in $\mathbf{X}$.
2. Scale variables in $\mathbf{X}$, e.g. scale to unit variance to give each variable equal weight in the PCA (Optional).
3. $n = 1$, $\mathbf{X}_n = \mathbf{X}$
4. First guess on score vector, $\mathbf{t}_n$ is first column in $\mathbf{X}_n$: $\mathbf{t}_n = \mathbf{X}_{n\bullet 1}$
5. Loading vector, $\mathbf{p}_n$ is calculated by projection of $\mathbf{X}_n'$ on $\mathbf{t}_n$: $\mathbf{p}_n = \mathbf{X}_n'\mathbf{t}_n \,/\, |\mathbf{X}_n'\mathbf{t}_n|$
6. Score vector, $\mathbf{t}_{n,\text{new}}$ is calculated by projection of $\mathbf{X}_n$ on $\mathbf{p}_n$: $\mathbf{t}_{n,\text{new}} = \mathbf{X}_n\mathbf{p}_n$
7. If $\mathbf{t}_{n,\text{new}} \neq \mathbf{t}_n$ then go to step 5, thus step 5 – 7 is the iterative loop which continues until convergence: $\mathbf{t}_{n,\text{new}} = \mathbf{t}_n$
8. Deflation of $\mathbf{X}_n$: $\mathbf{X}_{n+1} = \mathbf{X}_n - \mathbf{t}_n\mathbf{p}_n'$
9. $n = n + 1$
10. Go to step 4 as long as $n \leq N$
11. The residual, $\mathbf{E} = \mathbf{X}_{N+1}$

That is all. The PCA model is especially appealing due to the lack of *a priory* assumptions regarding the distribution of the variables and relation between them. The only assumption made is that data can be described reasonably well by a number of linear combinations determined by the data itself. In statistics, PCA is thought of as a useful descriptive tool among other more familiar tools such as means, medians, standard deviations and correlation coefficients. In that sense, PCA is an overview of data from which no general

[11] Alternatively singular value decomposition (SVD) may be used (Eckart & Young, 1936, Golub and Reinsch, 1970).

inference can be made. In chemometrics and applied science however PCA is additionally thought of as an exploratory hypothesis generating tool from which valuable general conclusions can be made, important variables found, outliers detected etc. (Esbensen, 2000). General conclusions should however be substantiated by proper validation, e.g. repeating the entire experiment and observe the same structure based on a new PCA. The beauty of this is that only the data speak with minimal interference from the researchers.

The following small illustrative example may serve as a numbers-on experience of NIPALS for those who have come this far and still are mystified by the magic of this explorative projection tool. This example takes you through PCA including various plotting options of scores and loadings for interpretation using a small example of simulated centred data. Centring the data serve as a magnifier focusing on sample differences and similarities, rather than the general level that usually is of minor interest. Four samples, s1 to s4, are investigated by two variables, V1 and V2. The data is presented in Fig. B1 as a data table, a matrix, as spectra, and as points in the variable space.

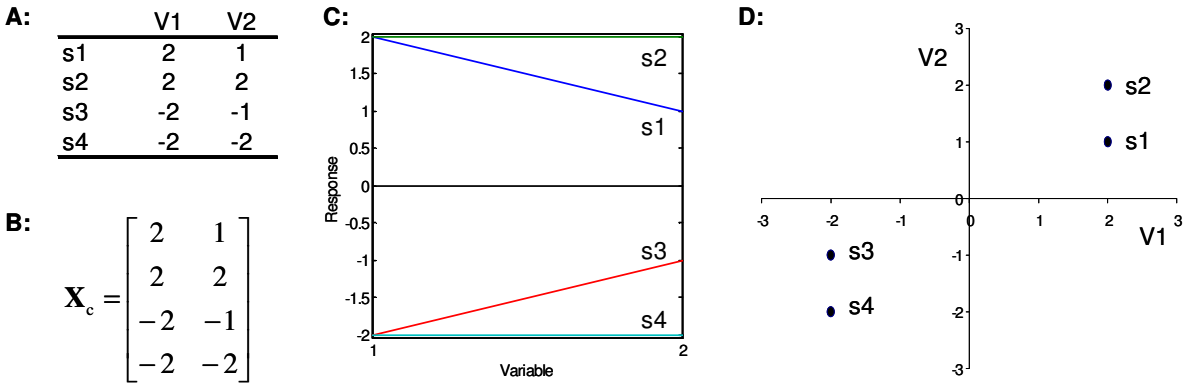


Figure B1: Simulated data. **A**: Data table with four samples in rows and two variables recorded in columns. **B**: The data in **A** represented by a matrix $\mathbf{X}_c$, c for centred. **C**: Data presented graphically as spectra. **D**: Each sample as a point in the Euclidian space spanned by variables V1 and V2.

From any representation of the data, A,B, C or D in Fig. B1 it is quite simple to see that sample 1 and 2 are quite different from sample 3 and 4 – and each two pairs are quite similar. Samples, s1 and s2, have high values in both V1 and V2 and vice versa for s3 and s4. It is easy to see in this example, but were there just a few more variables and a few more samples – the problem could turn very complex. The purpose here is to gain confidence in PCA by

seeing the expected and being able to follow the algebra numerically. We could just place a ruler somewhat close to the diagonal in Fig. B1D and state that it was surely the direction of the major variation and subsequently project the samples onto this line and read the scores. That is exactly what principal component analysis does.

Now we want to perform a principal component analysis. We want to establish the direction in the variable space spanning the major variation and the scores along the new variable. We start the PCA algorithm by choosing the first column of **X** as proxy score vector:

Step 4a: $$\mathbf{t}_1 = \begin{bmatrix} 2 \\ 2 \\ -2 \\ -2 \end{bmatrix}$$

Step 5a:

$$\mathbf{p}_1 = \begin{bmatrix} 2 & 2 & -2 & -2 \\ 1 & 2 & -1 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ -2 \\ -2 \end{bmatrix} \Big/ \left\| \begin{bmatrix} 2 & 2 & -2 & -2 \\ 1 & 2 & -1 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ -2 \\ -2 \end{bmatrix} \right\| = \begin{bmatrix} 16 \\ 12 \end{bmatrix} \Big/ \left\| \begin{bmatrix} 16 \\ 12 \end{bmatrix} \right\| = \begin{bmatrix} 0.8 \\ 0.6 \end{bmatrix}$$

This is a first suggestion of a direction in variable space. In fact, step 5a is a projection of $\mathbf{X}'$ onto $\mathbf{t}_1$ in the four dimensional sample space. It is not possible to depict this projection in four dimensions graphically at the present stage of human intellectual development. However projections in dimensions >3 is completely analogue to projections in dimensions ≤ 3 - in a mathematical sense.

Step 6a + 7a: $$\mathbf{t}_{1,\text{new}} = \begin{bmatrix} 2 & 1 \\ 2 & 2 \\ -2 & -1 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} 0.8 \\ 0.6 \end{bmatrix} = \begin{bmatrix} 2.2 \\ 2.8 \\ -2.2 \\ -2.8 \end{bmatrix} \neq \mathbf{t}_1$$

In Step 6a, **X** is projected on the suggested $\mathbf{p}_1$ direction in the variable space. Convergence criterion is not met in 7a, and $\mathbf{t}_{1,new}$ are now used as $\mathbf{t}_1$ when the algorithm iterates back to step 5:

Step 5b:

$$\mathbf{p}_1 = \begin{bmatrix} 2 & 2 & -2 & -2 \\ 1 & 2 & -1 & -2 \end{bmatrix} \begin{bmatrix} 2.2 \\ 2.8 \\ -2.2 \\ -2.8 \end{bmatrix} / \ldots\ldots\ldots = \begin{bmatrix} 20 \\ 15.6 \end{bmatrix} / \left\| \begin{bmatrix} 20 \\ 15.6 \end{bmatrix} \right\| = \begin{bmatrix} 0.7885 \\ 0.6150 \end{bmatrix}$$

is inserted into 6b and so forth until convergence at: $\mathbf{t}_1 = \begin{bmatrix} 2.192 \\ 2.807 \\ -2.192 \\ -2.807 \end{bmatrix}$ and $\mathbf{p}_1 = \begin{bmatrix} 0.7882 \\ 0.6154 \end{bmatrix}$, which is reached in just 3 iterations. In Fig. B2 it is now seen that the direction in the variable space indicated by $\mathbf{p}_1$ indeed appears to be the direction of the major variance in **X**. By projecting the samples onto this direction, the scores $\mathbf{t}_1$ can easily be verified graphically, this is why $\mathbf{t}_1 = \mathbf{X}\mathbf{p}_1$ in step 6 of the NIPALS Algorithm is called the projection of **X** on $\mathbf{p}_1$.

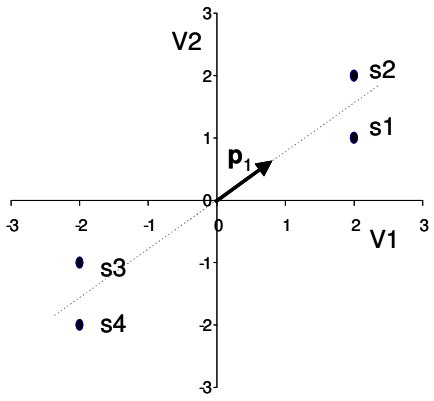


Figure B2: Unit vector $\mathbf{p}_1$ found by PCA. $\mathbf{p}_1$ is an intuitively reasonable direction spanning the major variation in data. By projecting each sample on to $\mathbf{p}_1$ their scores can be visualised.

In step 6 it is seen that $\mathbf{t}_1$ is a linear combination of **X**, and $\mathbf{p}_1$ and is a weighted sum of the original variables in which $\mathbf{p}_1$ contains the weights of

each original variable in the latent variable. The more weight a variable has, the more important it is in determining the direction in the original variable space. The loading vector $\mathbf{p}_1$ can also be considered a spectrum and the score values in $\mathbf{t}_1$ the amount of this spectrum in each of the samples. This component, $\mathbf{t}_1\mathbf{p}'_1$ is subtracted from X in the step 8:

Step 8a:

$$\mathbf{X}_2 = \mathbf{X}_1 - \mathbf{t}_1\mathbf{p}'_1 = \begin{bmatrix} 2 & 1 \\ 2 & 2 \\ -2 & -1 \\ -2 & -2 \end{bmatrix} - \begin{bmatrix} 2.192 \\ 2.807 \\ -2.192 \\ -2.807 \end{bmatrix} \begin{bmatrix} 0.7882 & 0.6154 \end{bmatrix} = \begin{bmatrix} 0.272 & -0.349 \\ -0.213 & 0.272 \\ -0.272 & 0.349 \\ 0.213 & -0.272 \end{bmatrix}$$

This step removes everything correlated to $\mathbf{t}_1$ in direction $\mathbf{p}_1$ and is thus called orthogonalisation of each element in $\mathbf{X}$ with respect to $\mathbf{p}_1$ or a deflation of $\mathbf{X}$. This can readily be seen in Fig. B3, where all points in space forms a line (a plane in higher dimension data). Now the algorithm returns to step 1:

Step 1a: $$\mathbf{t}_2 = \begin{bmatrix} 0.272 \\ -0.213 \\ -0.272 \\ 0.213 \end{bmatrix}$$

Step 2a:

$$\mathbf{p}_2 = \begin{bmatrix} 0.272 & -0.213 & -0.272 & 0.213 \\ -0.349 & 0.272 & 0.349 & -0.272 \end{bmatrix} \begin{bmatrix} 0.272 \\ -0.213 \\ -0.272 \\ 0.213 \end{bmatrix} / \ldots = \begin{bmatrix} 0.388 \\ -0.497 \end{bmatrix} / \left\| \begin{bmatrix} 0.388 \\ -0.497 \end{bmatrix} \right\| = \begin{bmatrix} 0.615 \\ -0.788 \end{bmatrix}$$

And so $\mathbf{t}_2$ becomes $\begin{bmatrix} 0.443 \\ -0.346 \\ -0.443 \\ 0.346 \end{bmatrix}$ in step 3a. As this is the last component, $\mathbf{p}_2$ and $\mathbf{t}_2$ are determined directly without extra iteration steps, logically because all point in $\mathbf{X}_2$ are exactly on a straight line orthogonal to $\mathbf{p}_1$ (Fig. B3).

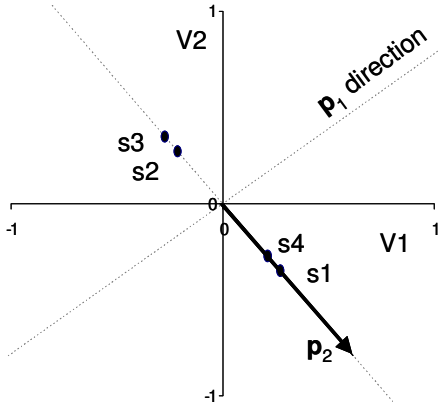


Figure B3: $\mathbf{X}_2$ plot (●) in the variable space. $\mathbf{X}$ has been collapsed along the $\mathbf{p}_1$ direction and reduced the dimensionality by 1. $\mathbf{X}_2$ is now 1-dimensional with no variation left along the first principal component.

Now, $\mathbf{T} = [\mathbf{t}_1\ \mathbf{t}_2]$ is the same as $\mathbf{X}$, just projected onto a two new latent variables given by $\mathbf{P} = [\mathbf{p}_1\ \mathbf{p}_2]$. The full model of X is thus:

$$\mathbf{X} = \mathbf{TP}' = \begin{bmatrix} 2.192 \\ 2.807 \\ -2.192 \\ -2.807 \end{bmatrix} \begin{bmatrix} 0.788 & 0.615 \end{bmatrix} + \begin{bmatrix} 0.443 \\ -0.346 \\ -0.443 \\ 0.346 \end{bmatrix} \begin{bmatrix} 0.615 & -0.788 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 2 \\ -2 & -1 \\ -2 & -2 \end{bmatrix}$$

Graphically the model can be viewed as weighted sums of components or spectra (Fig. B4).

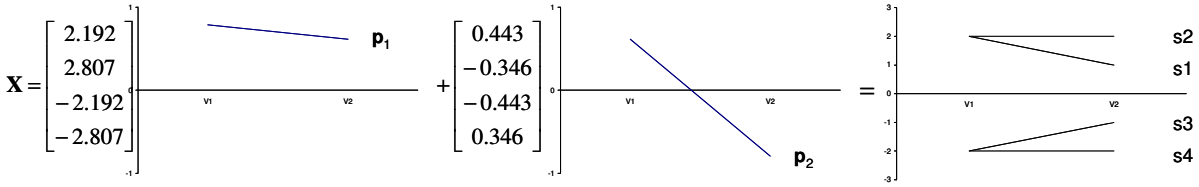


Figure B4: $\mathbf{X}$, decomposed as a linear combination of scores and loading spectra. Variable V1 and V2 are positively correlated in $\mathbf{p}_1$ and negatively correlated in $\mathbf{p}_2$.

Finally we will visualise the scores and loading in 2D plots. The score plot is a graphical representation of the samples in the PC space. The relevant PCs to inspect are those we critically choose as being significant. A number of ways to determine and validate the relevant PCs are described elsewhere (Esbensen, 2000) and will not be elaborated upon here. In this case we consider the full rank model with two principal components. When plotting loadings and scores simultaneously either next to each other or in the same

plot, it is important to plot on equal scales, as is done here. That is to get a correct mutual graphical interpretation of the data (Kroonenberg, 1997). This is not always the default option in software for multivariate data-analysis. Zooming on the plots for better visualisation of groupings and connections in samples and variables are of cause allowed, but caution should be exhibited not to over interpret the connections between scores and loadings.

Samples s1 and s2 are situated together along the first PC in the score plot in Fig. B5. They both have high scores on PC1, while s3 and s4 have equally low scores. The score plot thus depicts similarities among samples in **X**. PC1 explains 97.6% of the variation in **X**. PC2 explains the rest, 2.4% and is seen as s1 and s4 being slightly higher than zero, while s2 and s3 are slightly lower than zero.

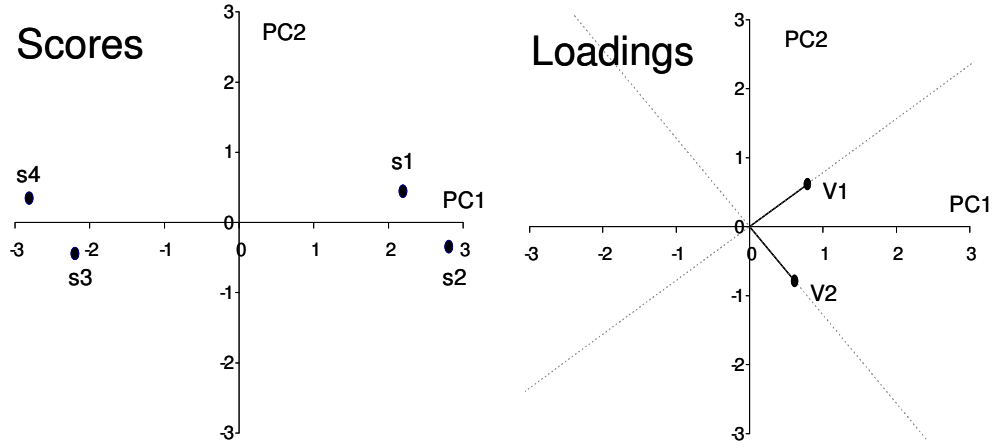


Figure B5: 2D scores and loadings plot of samples and variables in **X**. The figures confirm, what we already knew, s1 and s2 are similar and s3 and s4 are similar. The loadings plot shows, that both V1 and V2 contributes to the principal components.

The associated loadings plot depicts how the original variables span the principal components. It is seen that variable V1 and V2 have equal lengths of one. That is the case in higher order models too, however not necessarily visualised, but here the model is complete and we can see all dimensions. If we had chosen a one component model and considered PC2 random noise, loading vector V1 would be slightly longer than V2 along PC1. Only their projection on PC1 would be observed, i.e. 0.82 and 0.62. Generally high loadings far from the origin can be interpreted as important for the principal components observed. In this case V1 is slightly more important than V2 in explaining PC1.

Loadings are sometimes interpreted the same way as the scores above, i.e. variable loadings close together are co-varying while those far from each other are uncorrelated or independent. As we shall see – this could be wrong. In this case at least, V1 and V2 directions are orthogonal to each other, i.e. inner product is zero, and thus uncorrelated – if we follow the previous argument. By superimposing scores and loadings (Fig. B6), it is quite easy to see that, for each sample 2D-score, a projection can be made on the loading vectors, and the original variable values, x_{11}, x_{12}, x_{21} and x_{22} etc. read, as indicated on the plot. So effectively PCA is a rotation of data.

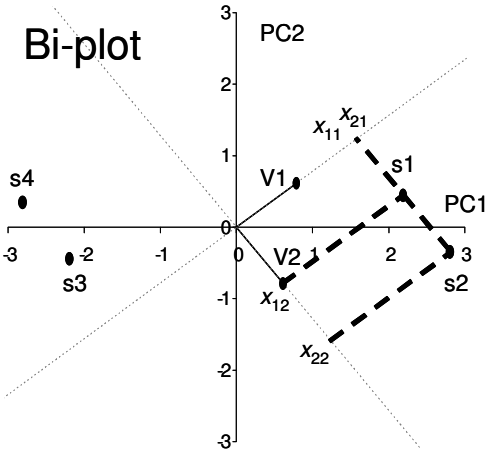


Figure B6: Scores and loadings superimposed in a 2D bi-plot with normalised **P** and all variance stored in **T**. Effectively a rotation of **X** with emphasis of successively describing as much of the variation as possible. The original **X** data can be read by projecting the samples onto the loadings.

And we also note that V1 and V2 are correlated in that high V1 is always associated with high V2 and vice versa. The correlation of V1 and V2 can also be calculated from the original data:

$$r_{V1V2} = \frac{\sum_{i=1}^{n}(v1_i - \overline{v1})(v2_i - \overline{v2})}{\sqrt{\sum_{i=1}^{n}(v1_i - \overline{v1})^2}\sqrt{\sum_{i=1}^{n}(v2_i - \overline{v2})^2}} = \frac{\sum_{i=1}^{n} v1_i v2_i}{\sqrt{\sum_{i=1}^{n} v1_i{}^2}\sqrt{\sum_{i=1}^{n} v2_i{}^2}} = \frac{\mathbf{v1}'\mathbf{v2}}{|\mathbf{v1}||\mathbf{v2}|} = \frac{\begin{bmatrix}2 & 2 & -2 & -2\end{bmatrix}\begin{bmatrix}1\\2\\-1\\-2\end{bmatrix}}{\sqrt{16}\sqrt{10}} = 0.949$$

But this correlation was not reflected graphically in the loading plots in Fig. B5 and B6, by loadings pointing in the same direction. This is however much better visualised if each vector in **T** is normalised with their lengths, thus making the $\mathbf{t}_n$ vectors orthonormal, hereafter called $\mathbf{u}_n$. n is the n'th principal component. The variability is then transferred to **P**. The lengths are effectively the eigenvalues or singular values, λ_1 and λ_2 in singular value decomposition (SVD)[12] (Eckart and Young, 1936).

$$\mathbf{X} = \lambda_1\mathbf{u}_1\mathbf{p}'_1 + \lambda_2\mathbf{u}_2\mathbf{p}'_2 \text{ and } \mathbf{t}_1 = \lambda_1\mathbf{u}_1, \quad \mathbf{t}_2 = \lambda_2\mathbf{u}_2$$

This means that we might as well scale the loadings with the singular values, rather than the singular vectors, $\mathbf{u}_i$. It is straight forward to calculate the singular values from the scores:

$$\lambda_1 = |\mathbf{t}_1| = \left\| \begin{bmatrix} 2.192 \\ 2.807 \\ -2.192 \\ -2.807 \end{bmatrix} \right\| = 5.04 \text{ and } \lambda_2 = |\mathbf{t}_2| = \left\| \begin{bmatrix} 0.443 \\ -0.346 \\ -0.443 \\ 0.346 \end{bmatrix} \right\| = 0.79$$

$$\text{Then } \mathbf{u}_1 = \frac{1}{\lambda_1}\mathbf{t}_1 = \frac{1}{5.04}\begin{bmatrix} 2.192 \\ 2.807 \\ -2.192 \\ -2.807 \end{bmatrix} = \begin{bmatrix} 0.435 \\ 0.557 \\ -0.435 \\ -0.557 \end{bmatrix}$$

$$\text{and } \mathbf{u}_2 = \frac{1}{0.79}\begin{bmatrix} 0.443 \\ -0.346 \\ -0.443 \\ 0.346 \end{bmatrix} = \begin{bmatrix} 0.557 \\ 0.435 \\ -0.557 \\ -0.435 \end{bmatrix}.$$

---

[12] SVD is an alternative to the NIPALS algorithm in which **X** is decomposed in singular vector matrices, **U** and **V**, with an additional diagonal matrix, **Λ**, containing the singular values, λ_n in decreasing order. The connection between PCA and SVD is straight forward: **X=UΛV′=TP′**, where **T=UΛ** and **P=V**. **Λ** contains the variance in each component and it can be shown that $\|\mathbf{X}\|^2 = \|\mathbf{\Lambda}\|^2$. λ_n^2 is thus the amount of variance explained by component n.

Now we can scale the loadings:

$$\mathbf{p}_{\lambda 1} = \lambda_1 \mathbf{p}_1 = 5.04 \begin{bmatrix} 0.7882 \\ 0.6154 \end{bmatrix} = \begin{bmatrix} 3.97 \\ 3.10 \end{bmatrix} \text{ and}$$

$$\mathbf{p}_{\lambda 2} = \lambda_2 \mathbf{p}_2 = 0.795 \begin{bmatrix} 0.6154 \\ -0.7882 \end{bmatrix} = \begin{bmatrix} 0.489 \\ -0.627 \end{bmatrix}$$

This construction is called principal component scaling and has the advantage, that the angles between loadings reflect their correlation, $r = \cos(\theta)$ (Kroonenberg, 1997). This was not evident from Fig. B5 and B6. The distance between samples in the plot is the standardised distance and does not reflect graphically with their closeness how similar they are with respect to the principal components as in Fig. B5 and B6. However, by projection of sample scores on the variable loading – similarities can be observed (Fig. B7). This representation is thus complementary to Fig. B6. Note that the loadings are still orthogonal, only their representations in the 2D-plot have changed.

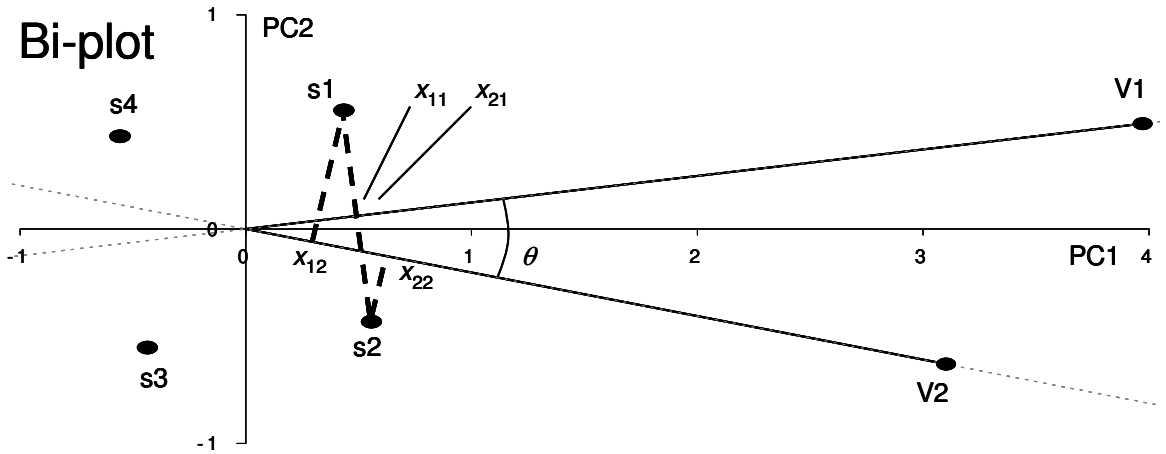


Figure B7: Principal component weighted bi-plot of scores and loadings in **X**. The cosine of the angle, θ, between the loadings is equal to the correlation between the variables, when the fit is perfect (Kroonenberg, 1997). The lengths of the loadings are proportional to their variance in the original data.

Often bi-plots are presented using symmetric scaling of scores and loadings in which the variability divided equally with $\mathbf{\Lambda}^{0.5}$ to **U** and **P**, respectively. The advantage is that scores and loadings are situated in the same plot as in Fig. B6 and B7 and sample similarities, variable similarities and evaluation

of original variable values by projection of scores onto loadings should ideally by available simultaneously (Fig. B8).

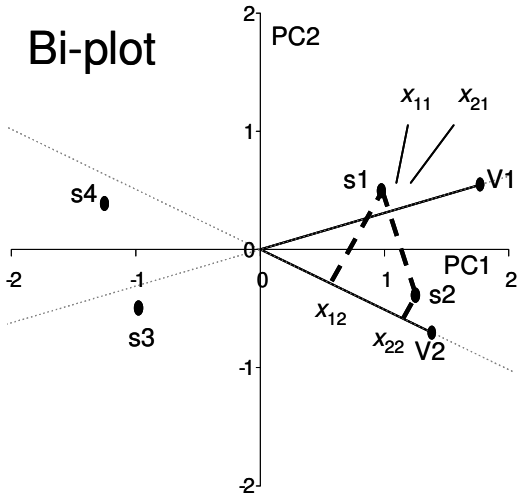


Figure B8: Symmetrically $\mathbf{\Lambda}^{0.5}$ scaled bi-plot of scores and loadings. Projections of sample points onto loading variables makes it possible to evaluate the relative magnitudes of variables (x_{11}, x_{12},..) for each sample. Samples does not appear as similar as in Fig. B5 and Fig. B6, neither does variables as in Fig. B7.

However the bi-plot with symmetric scaling in Fig. B8, although often used, is like sitting between two chairs. It is easy to interpret the samples as being quite different and the variables uncorrelated in Fig. B8. For a fast overview of data the above example points to an alternative representation of scores and loading rather than the conventional (Fig. B8). Two bi-plots with scaled scores (Fig. B7) and scaled loadings (Fig. B6) respectively should ideally be presented in order to both visualize sample similarities and differences as well as variable correlations and on the same time make projections of scores onto loadings possible for fast evaluation of original variables.

It is even tempting to suggest scaling both scores and loadings in the same plot. This might have the advantage of graphically pinpoint the sample to variable interaction in an instant (Fig. B9). Note however projections of sample scores onto variable loadings can only be assessed qualitatively, but powerfully indeed.

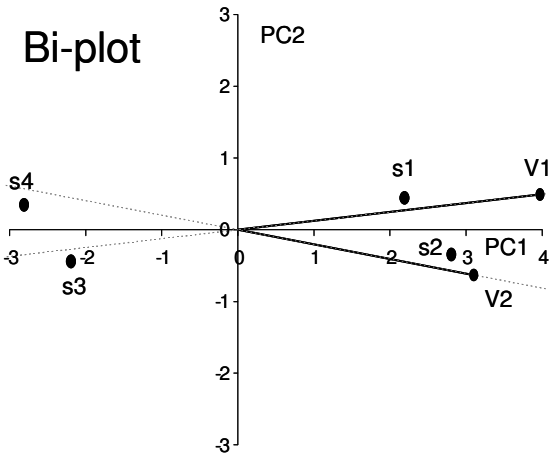


Figure B9: Bi-plot with both scores and loadings scaled with the SVD eigenvalues. This plot has the advantage of displaying both sample and variable relations simultaneously and thereby increase the interpretability. Projections of scores onto the loading is however not possible.

Note: If the data in this example (Fig. B1) were standardised, the picture would be more or less the same. Variables would have equal weight and loadings would become equally important with lengths one. Loadings and scores appear in the same space in bi-plots which could be an advantage when interpreting data. However, standardising does not solve the risk of over- or under-interpreting scores/loadings and bi-plots as outlined above.

Paper I

Protein heterogeneity in wheat lots using single-seed NIT — A Theory of Sampling (TOS) break-down of all sampling and analytical errors

E. Tønning, L. Nørgaard, S.B. Engelsen, L. Pedersen, K.H. Esbensen

Chemometrics and Intelligent Laboratory Systems, 84, 152-162 (2006).

Paper II

Bulk quality diversification of organic wheat by single-kernel near-infrared (SKNIR) sorting

E. Tønning, A.K. Thybo, L. Pedersen, L. Munck, Å. Hansen,
S.B. Engelsen, L. Nørgaard

Cereal Chemistry, submitted.

Paper III

Stepwise multivariate prediction of wheat flour functionality and bread quality

E. Tønning, A.K. Thybo, L. Pedersen, L. Munck, F. van den Berg, S.B. Engelsen, L. Nørgaard

Cereal Chemistry, submitted.

Paper IV

A novel improved method for analysis of 2D diffusion–relaxation data —2D PARAFAC-Laplace decomposition

E. Tønning, D. Polders, P.T. Callaghan, S.B. Engelsen

Journal of Magnetic Resonance, 188, 10-23 (2007).

Available online at www.sciencedirect.com
ScienceDirect
Chemometrics and Intelligent Laboratory Systems 84 (2006) 142–152
Chemometrics and intelligent laboratory systems
www.elsevier.com/locate/chemolab

Protein heterogeneity in wheat lots using single-seed NIT — A Theory of Sampling (TOS) breakdown of all sampling and analytical errors

Erik Tønning [a,*], Lars Nørgaard [a], Søren B. Engelsen [a], Lene Pedersen [b], Kim H. Esbensen [c]

[a] *Quality and Technology, Department of Food Science, The Royal Veterinary and Agricultural University (KVL), Centre for Advanced Food Studies, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark*

[b] *Plant Food Science, Department of Food Science, The Danish Institute of Agricultural Sciences (DIAS), Centre for Advanced Food Studies, DK-5792 Årslev, Denmark*

[c] *ACABS (Applied Chemometrics, Analytical Chemistry, Acoustic Chemometrics, Applied Biotechnology, Bioenergy and Sampling research group), Aalborg University Esbjerg (AUE),DK-6700 Esbjerg, Denmark*

Received 6 February 2006; received in revised form 2 May 2006; accepted 3 May 2006
Available online 11 July 2006

Abstract

An in-depth heterogeneity analysis of wheat lots from varying field experiments with respect to protein concentration was conducted in order to quantify and compare both sampling and analytical errors as defined by the Theory of Sampling (TOS). Thirty wheat samples of forty-two seeds were extracted from three different wheat lots. Half of these were extracted using a non-optimal spoon (grab sampling) and the other half were extracted using a riffle splitter. Ten additional samples of forty-two seeds were extracted using a riffle splitter from ten different wheat lots. The protein content of every single-seed was determined by Near-Infrared Transmission (NIT) spectroscopy based on a multivariate calibration to Kjeldahl with Partial Least Squares Regression (PLS-R). The effect of orientation and number of replicate measurements of the individual seeds in the NIT beam was investigated for minimizing the analytical error as well the resulting time requirements. Remarkably, the best prediction model with respect to Root Mean Squared Error of Cross Validation (RMSECV) was obtained by only recording three replicate NIT spectra of the seeds in only one specific orientation.

The variance of the Global Estimation Error (GEE) of both the riffle splitter and the spoon sampling processes was estimated as well as its components, the Fundamental Sampling Error (FSE), the Grouping and Segregation Error (GSE), the Incorrect Sampling Errors (ISE) and the Total Analytical Error (TAE). The bias induced by non-probabilistic spoon extraction was also estimated. The GEE variance of the spoon extractions was seventy percent higher than that of the riffle split samples. The sampling variances of FSE, GSE and ISE were all of the same order of magnitude, each approximately ten times higher than the TAE variance. The squared bias of the spoon sampling was approximately twice the magnitude of the sampling variances and thus contributed significantly to the representativity score. Spoon sampling representativity was three times the size of that for the riffle splitter. Order of magnitude estimates of the Constitutional Heterogeneity (CH_L) as well as the Distributional Heterogeneity (DH_L) for a forty-two seed riffle split sample was derived. In this investigation the fundamental concepts of TOS have been investigated and estimates of all sampling and analytical errors have been presented for a specific zero-dimensional composite material. The ability of TOS for quantifying and evaluating the various error contributions to the overall estimation of the protein concentration was confirmed.

Keywords: Theory of Sampling (TOS); Heterogeneity; Wheat; Sampling error; Analytical error; Single-seed NIT; Single-kernel analysis; Representative sampling; Uniform materials

1. Introduction

When scientists, technicians or authorities measure any property, spectrum or concentration of critical components in a set of samples, the objective is to obtain results which are approximately correct — meaning both accurate and reproducible, seeking representativity of the analytical sample with

* Corresponding author. Tel.: +45 35283264; fax: +45 35283245.
E-mail address: ert@kvl.dk (E. Tønning).
URL's: http://www.models.kvl.dk (E. Tønning), http://www.models.kvl.dk (L. Nørgaard), http://www.models.kvl.dk (S.B. Engelsen), http://www.agrsci.dk (L. Pedersen), http://www.acabs.dk (K.H. Esbensen).

0169-7439/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemolab.2006.05.007

respect to the lot from which it was extracted. Thus, in the sampling process, from lot to analytical volume, a mass reduction that may exceed 1:100.000 comes into focus [1–7]. It is not possible to compensate a weak primary sampling under any circumstances.

For quality assurance and error estimation, analysis is often performed in replicates. Barring the circumstances where assurance may be constrained by lack of available material, costs, analytical safety or time, the practice of replication cannot be emphasized enough. In case of largely diverging or unexpected results, experiments are sensibly repeated to investigate the suspicious variation. However, this practice may also cause biased results and underestimation of the true variance: This occurs when the unexpected and apparently unexplainable results are discarded in favour of the expected, less diverging or explainable results. Apart from pure operational errors that should ideally be the sole reason for discarding results, the origin of anomalous results may often be explained by the fact that the lot and samples derived from it are indeed more heterogeneous than assumed and that consequently the countermeasures taken, i.e. replication of experiments or mixing, are inadequate.

The absence of homogeneity in the physical world is the justification of Pierre Gy's Theory of Sampling (TOS) [1–7]. All properties of natural systems and samples are to some degree always spatially unevenly distributed. Homogeneity is a scale dependent feature and we tend to categorize systems with apparent low heterogeneity such as for example grains, flour, suspensions, fluids, sand on the beach, soil or lots of (apparently) identical products as homogeneous. In some instances the heterogeneity scale is many orders of magnitude smaller than the scale of the sampling instrument in which cases heterogeneity may not be the main source of variation [8]. However, as this study will demonstrate, assuming homogeneity without preceding investigation or theoretical considerations should be avoided — with the exception of solutions in thermodynamic equilibrium [9].

TOS defines a sampling process of a given lot to be representative, and thus unbiased, when all fragments or groups of fragments have an equal probability of ending up in the sample. This fundamental sampling principle [1] can always be applied when the lot is zero-dimensional; a zero-dimensional lot is by definition a three-dimensional collection of material in which all parts of the material are equally accessible — possibly with a little work, e.g. a pile of grain, a bag of soil or lots of wheat seeds as investigated in the present case. In TOS the fundamental sampling principle is referred to as *correct sampling*, which is a technical term used to signify that all appropriate principles contained in TOS leading to a demonstrable representative sampling process have been invoked. Strictly representative sampling may sometimes not be possible when the lot by size or higher order dimensionality is appropriate to handle as a zero or one dimensional lot [10], e.g. a landfill site (2-D), a mining site (3-D) or an invaluable historical relic (3-D). 2-D and 3-D are indeed special cases with problem-dependent solutions. For more on lot dimensionality refer to Petersen et al. [11].

This study aimed to perform an in-depth heterogeneity analysis to estimate and compare the various error contributions associated with the sampling of apparently homogeneous lots of wheat grain. This type of material is usually referred to as "uniform", and is commonly considered to be significantly less heterogeneous than what merits thorough sampling considerations. Mass reduction in the form of representative riffle splitting is compared to grab sampling using a spoon to extract samples. The unique situation with grain lots compared to most other materials is that it is possible to measure the protein grade of every single seed using near-infrared spectroscopy as well as the mass due to their manageable size. Others have used single-seed near-infrared spectroscopy to classify [12,13] and determine protein content [14–19] as well as other characteristics of wheat [14,19–21]. Delwiche and Hruschka [18] presented single-seed measurements as a method for determining the bulk protein concentration using near-infrared reflectance spectra from 300 single seeds. The method had a reproducibility comparable to the usual bulk protein determination of whole grains using near-infrared transmittance NIT with the advantageous added ability of retaining information about the variability [17,18], i.e. heterogeneity of the lot.

The specific feature of the present study, compared to nearly all other sampling work, is that the composition of each grain is available through single-seed NIT analysis — each analytical result equating the composition of each *fragment* in the TOS parlance. This will make it possible to directly calculate the compositional heterogeneity, a feature which is otherwise very nearly always only estimated by indirect, approximate means [1–6,8], which will allow the study to present a complete breakdown of all sampling errors involved in grain characterisation, including an independent optimization of the orientation of the single seeds in the NIT beam, leading to an estimate of the analytical error as well as a quantification of the compositional as well as the distributional heterogeneity of typical grain lots. The applicability of TOS for 0-dimensional lots is investigated in the present study and the implication for other composite particulate samples should be readily appreciated.

2. Theory

Sampling is representative mass reduction of the lot, i.e. the material of interest, to allow the analytical procedure to produce the data output desired. However, the data recorded from the analysis are only estimates of the true properties of the lot [11]. For estimation of the grade, a_L, i.e. the true mass proportion of the critical component in the lot, L, the relative errors of the analytical result, a_R, are defined by TOS:

The global estimation error (GEE):

$$\mathrm{GEE} = \frac{a_R - a_L}{a_L}. \tag{1}$$

The total sampling error (TSE):

$$\mathrm{TSE} = \frac{a_S - a_L}{a_L}. \tag{2}$$

The total analytical error (TAE):

$$\mathrm{TAE} = \frac{a_R - a_S}{a_L}. \tag{3}$$

where a_S is the true unknown grade of the critical component in the sample and:

$$\mathrm{GEE} = \mathrm{TSE} + \mathrm{TAE}. \quad (4)$$

TOS further specifies that TSE is composed of three additive errors: the fundamental sampling error (FSE), the grouping and segregation error (GSE) and the incorrect sampling errors (ISE):

$$\mathrm{TSE} = \mathrm{FSE} + \mathrm{GSE} + \mathrm{ISE}. \quad (5)$$

FSE is an inevitable error due to the variation in grade, mass, size or shape of the individual fragments, *F*, in the lot. TSE is fundamentally limited by FSE, and is as such a material constant that can only be reduced by crushing the lot into even smaller particles before the sampling process. GSE is the error arising from fragments not being randomly distributed in the lot, but instead are spatially grouped to various degrees due to poor mixing or forces making the fragments stick together. In particulate material, gravitational segregation very nearly always contributes to the internal sorting of particles by size, density and shape during transport, handling, storage and potential mixing [1,11]. GSE can be reduced by mixing the lot, by extracting more and smaller increments from the lot or by introducing crushing and dissolution steps in the sampling process. Dutch legislation on the reuse of soil has recently implemented such steps in their sampling strategy. In order to obtain cost effective estimates of the degree of contamination, single representative analytical results are obtained from composite samples that undergoes both mixing, splitting and crushing steps [7]. FSE and GSE are both errors that are defined as correct sampling errors (CSE) [3]. Extracting samples in a non-probabilistic way, e.g. by taking the sample from an easily accessible site or by using tools and procedures that generate failures in delimitation, extraction and handling of the sample, generates errors that overall are defined as incorrect sampling errors (ISE). We refer to Gy [1,3] and Petersen et al. [11] for a thorough description of these errors. ISE are purely circumstantial and may be small or large depending on the material and the sampling process; ISE do not follow any statistical distribution. ISE is the source of biased analytical results that may or may *not* have a large variation when repeated. Making replicate measurements to qualify the quality of any measurement by stating the variance or standard deviation *alone* is thus *inadequate* if not accompanied by a full documentation of the sampling process, because using a fully representative sampling process is the only way to avoid an uncontrolled, indeed impossible-to-estimate, sampling bias [1–6,8,11].

The implication of the relationship between these sampling errors is that it is only possible to arrive at a sampling error of comparable magnitude to FSE by extracting the sample fragment by fragment in a totally random fashion. This is, of course, not practically implementable for most materials. However, samples comprising of multiple randomly selected increments, *I*, of groups of fragments is usually possible; this process is called *composite sampling*. Extracting samples, increment by increment, is what makes GSEs inevitable with an impact dependent on the degree of mixing and the number of increments in the final sample. Representative sampling can thus be said to be about eliminating ISE, while simultaneously reducing GSE and TAE maximally — to achieve a replication variation in GEE of an acceptable level determined by the context. Within this TOS context, selecting an appropriate sample size is very much dependent upon the material heterogeneity in addition to the sampling process itself. TOS is adamant in arguing that choosing sample size is impossible without an appropriate heterogeneity analysis of the material in question [1–4]. This aspect has been operationalised recently in a very hard-headed, practical context [11].

Assessing the essential relationships between the individual sampling errors can be stated in a statistical fashion by their means and variances. The mean error or the bias of replicate measurements is:

$$m(\mathrm{GEE}) = m(\mathrm{FSE}) + m(\mathrm{GSE}) + m(\mathrm{ISE}) + m(\mathrm{TAE}), \quad (6)$$

while the variance of GEE is a sum of variances:

$$\sigma^2(\mathrm{GEE}) = \sigma^2(\mathrm{FSE}) + \sigma^2(\mathrm{GSE}) + \sigma^2(\mathrm{ISE}) + \sigma^2(\mathrm{TAE}). \quad (7)$$

Pierre Gy [3] in a particularly prescient mood, redefines the concepts of accuracy and reproducibility by stating that a sample is *accurate* when:

$$m(\mathrm{ISE}) = 0 \quad (8)$$

leading to:

$$m(\mathrm{TSE}) \approx 0, \quad (9)$$

thus making accuracy a property of the sampling process, i.e. the rule of uniform selection probability must be obeyed to state that a sample is unbiased (fundamental sampling principle).

Reproducibility, sometimes denoted precision, is based on a user-defined threshold, σ_0^2, which the variance of TSE must not exceed:

$$\sigma^2(\mathrm{TSE}) = \sigma^2(\mathrm{FSE}) + \sigma^2(\mathrm{GSE}) + \sigma^2(\mathrm{ISE}) <= \sigma_0^2. \quad (10)$$

In addition, Gy also defines a new score, the representativity as the sum of the variance and the squared bias of TSE. Thus, for a sample to be representative it must be both accurate and reproducible. For this, the user-defined threshold of representativity, r_0^2, must not be exceeded:

$$r^2(\mathrm{TSE}) = m^2(\mathrm{TSE}) + \sigma^2(\mathrm{TSE}) <= r_0^2. \quad (11)$$

The representativity thus defined is a useful score for ranking different sampling processes and equipment [22]. By introducing the representativity as a score that can be either under or above the user-defined threshold, rather than a nebulous principle, Gy [3] allows for certain flexibility with respect to the accepted bias. This is to state that it is acceptable to use a sampling protocol which is not strictly representative as long as the level of bias is known and

Table 1
Summary of riffle split the wheat lots with an overview of their characteristics including estimates of the true protein concentrations, constitutional heterogeneity and the fundamental error associated with protein determination of the lot using one or an average of forty-two randomly selected seeds

ID	Cultivar	Year	Growth place in Denmark	Lot mass M_L [g]	# seeds in FSE calc.	Av. seed mass M_{i*} [mg]	Protein conc a_L [%dm]	FSE: sample=one seed				FSE: sample=forty-two seeds		
								s^2(FSE) [$\times 10^{-3}$]	s(FSE) [%]	s [%dm]	CH_L [$\times 10^{-3}$]	s^2(FSE) [$\times 10^{-5}$]	s(FSE) [%]	s [%dm]
W01	Pentium	2002	Norsminde	4686	203 (7)	48.0	11.1	12.6	11%	1.2	11.4	27.1	1.6%	0.2
W02	Ritmo	2002	Ruballegaard	2833	203 (7)	39.6	10.6	19.2	14%	1.5	20.7	49.3	2.2%	0.2
W03	Claire	2002	Ruballegaard	2631	204 (6)	40.0	9.9	14.8	12%	1.2	15.0	35.7	1.9%	0.2
W04	Bussard	2002	Harlev	3169	39 (3)	40.6	10.4	11.3	11%	1.1	8.8	21.0	1.4%	0.2
W05	Galatea	2000	Abildgaard	7824	41 (1)	52.9	11.2	24.5	16%	1.8	28.0	66.7	2.6%	0.3
W06	Batis	2002	Kiel (G)	2876	39 (3)	43.4	13.1	21.7	15%	1.9	21.0	50.0	2.2%	0.3
W07	Bussard	2002	Kiel (G)	3726	41 (1)	42.6	12.5	25.9	16%	2.0	27.2	64.8	2.5%	0.3
W08	Renan	2002	Kiel (G)	3818	38 (4)	46.5	10.8	21.3	15%	1.6	24.7	58.8	2.4%	0.3
W09	Pentium	2003	Aarslev	656	38 (4)	43.1	11.8	44.0	21%	2.5	36.4	86.7	2.9%	0.3
S10	Vinjett	2003	Aarslev	671	38 (4)	41.0	11.4	39.5	20%	2.3	46.6	111.0	3.3%	0.4
S11	Leguan	2002	Foulum	612	39 (3)	34.1	14.1	18.6	14%	1.9	21.7	51.7	2.3%	0.3
S12	Leguan	2002	Foulum	584	42 (0)	34.4	12.5	13.6	12%	1.5	13.5	32.1	1.8%	0.2
S13	Leguan	2002	Foulum	467	37 (5)	33.9	15.8	19.2	14%	2.2	16.4	39.0	2.0%	0.3
			(G)=Germany		(# outliers)		TAE:	*0.6–1.6*	*2.5–4.0%*	*0.4*		*1.5–3.9*	*0.4–0.6%*	*0.06*

Boldfaced lots were also used for spoon sampling. For comparison the variance of total analytical error is indicated on the bottom line and is not subtracted from the values in the table, as indicated in Eq. (14).

within an accepted threshold. However, bias is only very rarely known — and why bother too much about this when representative sampling equipment actually does exist for a wide variety of purposes.

Riffle splitters and similar dividers are specifically designed to eliminate ISE and reduce GSE by insuring a uniform selection probability throughout the mass reduction. Their design and performance vary considerably, however. This was studied by Petersen et al. in a comprehensive overview [22], in which seventeen types of mass reduction equipment were investigated and compared with respect to accuracy, reproducibility, representativity as well as practical applicability by sampling a composite particulate material of known composition. It was shown that many of them are by design not working properly, actually introducing bias and excessive variation into the sampling process. All of the tested splitters and related dividers were orders of magnitude superior to various spoon and shovel methods investigated with respect to accuracy, reproducibility and for some also ease of use. Other studies resulted in similar findings with other particulate materials [23,24].

In commercial, industrial and academic laboratories, the materials under investigation often vary tremendously in total mass and physical properties, thus affecting which scale and mass reduction technique is most appropriate. For instance, most of the above mentioned dividers are all but useless if the materials are wet or sticky, or when the particle aggregates do not have reasonable flow properties. In order to perform sound mass reduction of such a variety of samples laboratories should ideally have a set of dividers and techniques, in multiple scales, to suit all appropriate sampling needs.

2.1. Calculations of TOS parameters

From the Root Mean Squared Error of Cross Validation (RMSECV) of a Partial Least Squares Regression (PLS-R) protein calibration it is possible to determine the variance of TAE for a sample composed of N seeds.:

$$s^2(\mathrm{TAE}) = \frac{\mathrm{RMSECV}^2}{a_L^2 \cdot N}. \quad (12)$$

Note that throughout this paper variances and standard deviations are presented as: s^2(relative error) and s(relative error) e.g. s^2(TAE) and s(TAE), which is equivalent to coefficient of variance, CV^2 and CV and relative standard deviation, RSD^2 and RSD.

Being able to analyse individual grains makes it is possible to estimate the constitutional heterogeneity with respect to the protein grade, CH_L, which is a material constant:

$$CH_L = s^2(h_i), \quad (13)$$

where h_i is the heterogeneity contribution of the individual seeds, i.e. the errors weighted by the individual masses, M_i, to the average seed mass, M_i*:

$$h_i = \frac{a_i - a_L}{a_L} \cdot \frac{M_i}{M_{i*}}. \quad (14)$$

From CH_L the variance of FSE for a defined increment size with N seeds can be deduced from Gy [1]:

$$s^2(\mathrm{FSE}) = \left(\frac{1}{M_S} - \frac{1}{M_L}\right) CH_L M_{i*} - s^2(\mathrm{TAE}) = \left(\frac{1}{NM_{i*}} - \frac{1}{M_L}\right) \times CH_L M_{i*} - s^2(\mathrm{TAE}) \approx \frac{CH_L}{N} - s^2(\mathrm{TAE}) \quad (14)$$

Having established both TAE and FSE it is possible to estimate the variance of GSE by subtraction rearranging Eq. (7):

$$s^2(\mathrm{GSE}) = s^2(\mathrm{GEE_{split}}) - s^2(\mathrm{TAE}) - s^2(\mathrm{FSE}) \quad (15)$$

The variance of GEE_{split} is easily calculated from replicate splitting of the sample. Likewise, the variance of ISE can be estimated from:

$$s^2(\mathrm{ISE}) = s^2(\mathrm{GEE}_{\mathrm{spoon}}) - s^2(\mathrm{FSE}) - s^2(\mathrm{GSE}) - s^2(\mathrm{TSE}) \qquad (16)$$

In the same way the biases of the individual sampling strategies can be calculated using Eq. (6).

3. Experimental

3.1. Material

From four different field trials in Denmark and in Germany, thirteen wheat lots with masses from 467 g to 7824 g were chosen based on availability and expected mutual variation in both physical and chemical properties as well as variations in constitutional heterogeneity. The lots represented seven winter and two spring wheat cultivars, with varying growth conditions, which will not be discussed further here for the present purpose (Table 1).

3.2. Sampling

Five spoon samples, each consisting of forty-two seeds, were obtained from each of the lots W01, W02 and W03. Each extraction was initiated by stirring the lot thoroughly followed by shovelling forty-two seeds into the spoon. Five split samples, also of forty-two seeds, were subsequently obtained from each of the same three lots above. Single samples of the same size of the remaining ten lots were likewise extracted. All of these latter samples were extracted using a riffle splitter from Rationel Kornservice with 18 chutes of width 16 mm each. The splitter was ranked among the better acceptable sampling devices in the survey of mass reduction hardware [22]. This instrument splits the original sample in two representative sub-samples of equal mass. One was chosen by random selection, flipping a coin, and subsequently split again. This procedure was repeated until the mass of the last randomly chosen half was between 10 and 20 g; the last split was thus performed on a minimum of 20 g of sample (Fig. 1A). After this point the riffle splitter was no longer able to perform its function in a proper way, as the subsequent splits would be too coarse in relation to the sample and fragment (seed) sizes. To continuously ensure uniform selection probability the resulting 10–20 g subsample consisting of approximately 250–500 seeds were laid out on a line in random order. Starting from a random position among the first six seeds a 1:5 split was performed by systematically selecting every sixth seed, disregarding broken seeds (Fig. 1B). From the now countable number of seeds still in line, forty-two [25] seeds were randomly chosen. The resulting forty-two seeds-size sample was thus effectively obtained by randomly extracting it fragment by fragment, although still affected by the internal grouping and segregation in the original lot as well as throughout the sampling process. In order to retain identity throughout the subsequent measurements, every single seed was stored separately in marked microtiter plates at 15–20 °C. In all, forty samples of forty-two individual marked seeds (1680 seeds) were obtained, i.e. five spoon samples and five split samples of W01, W02 and W03 respectively and one split sample of W04 to S13 respectively. With respect to the lot masses (Table 1), samples of forty-two seeds represent mass reductions in the range from 1:3500 to 1:330.

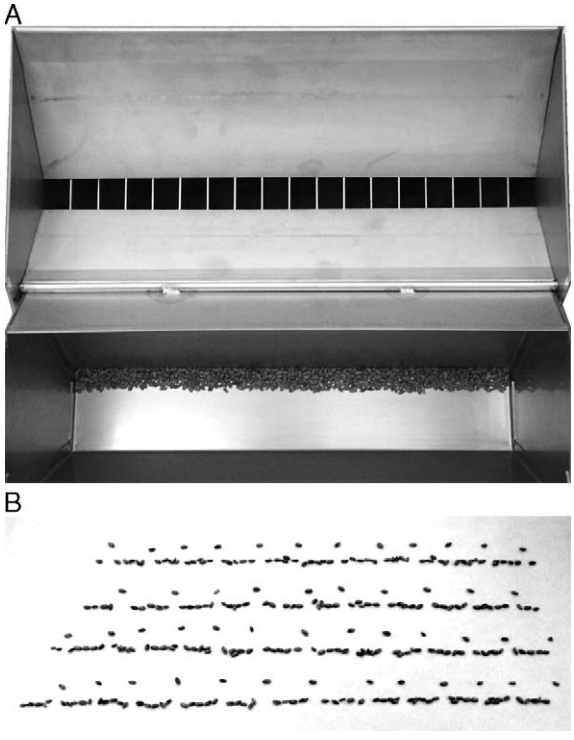


Fig. 1. A: 20 g wheat seed sample ready for the last mass reduction split. This is the smallest mass split by the Rational Kornservice riffle splitter. B: A 10 g sample from the riffle splitter equating 290 seeds arranged in a linear array. Every 6th seed is picked starting randomly from seed 1 to 6, in this case No. 2. A total of 48 seeds were picked and reduced to 42 by random selection and stored separately in order to retain identity.

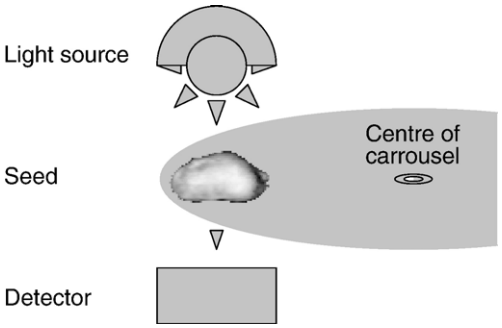


Fig. 2. Sample presentation in the Infratec 1255 Food and Feed Analyzer fitted with the single seed autosampler cassette. The seed is oriented here with the elongated furrow down towards the detector and the germ pointing to the centre of the carrousel.

For calibration purposes a subset of seven seeds from each of the samples of W01 to S13 were chosen by random selection. The resulting ninety-one single seeds thus represented the major variations in the material composition. The calibration set was used for protein calibration and optimization of orientation in the NIT beam (below).

For reference (see next paragraph) a subset of four seeds was randomly selected from W01, W03, W05 and S10.

3.3. NIT spectroscopy

The near-infrared spectra were recorded on an Infratec 1255 Food and Feed Analyzer from FossTecator, Höganäs, Sweden, mounted with a single seed autosampling cassette with slots for twenty-three single seeds. The time consumption for measuring one spectrum of each of the twenty-three seeds was approximately 90 s. Nineteen seeds were measured in every single run along with the four reference seeds. The position of the reference seeds were changed randomly between the slots with every change of seeds in order to asses the performance of the individual slots as well as the effect of time, since measurements were performed over a period of eight months. For the ninety-one calibration and 4 reference seeds near-infrared transmission spectra from 850 nm to 1050 nm were recorded in triplicate in eight different orientations of the seeds in the cassette, i.e. with the elongated furrow up, down, to the left or to the right and the germ either towards the centre or towards the edge of the cassette. In Fig. 2, a schematic representation of the experimental setup relative to the NIT beam is shown with the seed oriented with the elongated furrow down and the germ towards the centre of the cassette.

This orientation was found to be optimal (see Results and discussion), thus all other 1585 seeds as well as the reference seeds were measured in triplicate in this orientation. Between each of the triplicate measurements the seeds were positioned again, so that the position of the individual seed in the slot was not exactly identical in the three spectral recordings. This introduces a realistic "presentation variance".

3.4. Mass, dry matter content and protein

The masses of all individual in-situ seeds were recorded on a Mettler/Toledo scale, type AB204. The ninety-one calibration seeds were then crushed and dried at 130 °C for 2 h and their masses were recorded again for determination of dry matter content. The nitrogen content was determined by a modified Kjeldahl method according to AACC Method 46-12 previously reported by Pedersen et al. [15]. The replication error of the method was shown to be 0.16% dry matter protein and thus a relative error of 1–2%, since the protein %dm is in the range from 9.9 to 15.8.

3.5. Calibration

In order to evaluate the effect of the different orientations of the seeds in the near-infrared light path, NIT spectra were selected individually or collectively and averaged in various ways and correlated to reference protein content with partial least squares regression (PLS-R) [26,27]. The selected and averaged spectra were mean centred and scatter-corrected before calibration by applying a combination of second derivative (Savitzky–Golay, 2nd order polynomial) followed by multiplicative scatter correction (MSC) which has been found to work well for the single-seed protein system [15,16]. For convenience, interval partial least squares regression (*i*PLS) [28] in MATLAB version 6.5 (The MatWorks, Inc., Natwick, MA) was used to calculate the numerous models by constructing a consecutive series of the selected spectra before analysis. The models were cross validated using leave one out (LOO) validation and their performance with regard to the root mean squared error of cross validation (RMSECV) could easily be assessed. Outlying recordings from six seeds, three winter and three spring cultivars from different lots, were discarded in the calibration process due to unparalleled behaviour in the models. Three of the outlying seeds were either comparatively small or odd-shaped, while the other three did not have apparent deficiencies judged from digital pictures taken before crushing, however they were all recorded in the last bin (no. 23) of the autosampler cassette.

3.6. Prediction

Protein concentration predictions of all seeds were performed using Camo Unscrambler 8.0 on the basis of the optimal model, i.e. with the lowest RMSECV. Eighty of 1589 recordings with prediction deviations (based on both the leverage and the residual of the new sample) above 0.84, i.e. above the range found for the calibrations seeds (0.18–0.82), were discarded as outliers. The main reasons for large deviations were systematic errors due to occasional misalignment of the last bin in the autosampler cassette (30 recordings) as well as odd-shaped seeds (31 recordings), as were found for the calibration outliers, while 19 recordings had to be discarded due to operational errors. Delwiche [16] also found small seeds most difficult to measure — with higher prediction errors if measurement and predictions were repeated.

3.7. Calculation of TOS parameters

In this investigation only five replications were made of each of the lots W01, W02 and W03 (Table 1). In order to assess the order of magnitude of the above mentioned variances and biases associated with grain sampling, variances and squared biases were averaged before presentation in Fig. 5. An average of the distributional heterogeneity, DH_L, associated with the splitting technique down to forty-two seeds was also calculated using Eqs. (13) and (14) keeping in mind that index, "i", now refers to an increment of forty-two seeds.

4. Results and discussion

The first goal was to establish a single-seed protein NIT calibration meant to span the entire variation to be encountered in thirteen different wheat lots. Although the calibration is only based on seven randomly selected seeds from each of the samples originating from these lots, it does show a remarkable robustness, as was also the case in previous studies on single-seed investigations such as Delwiche [16]. In a comparison with five other pre-treatment strategies including no pre-treatment and extended inverted signal correction (EISC) [29], Pedersen et al. [15] showed that fewer PLS components and a lower RMSECV and RMSEP was found by pre-treating averaged single-seed NIT

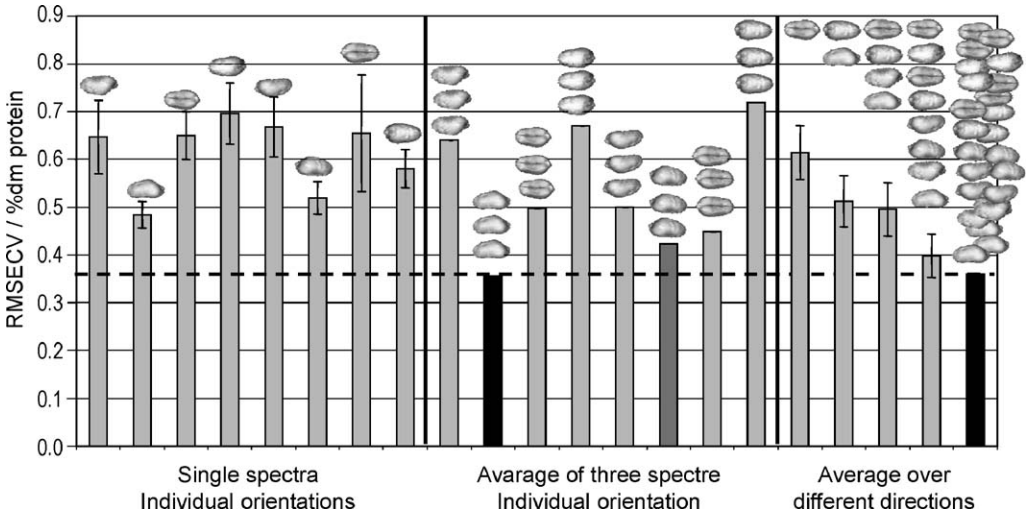


Fig. 3. RMSECV of 21 individual PLS-R calibrations as presented by i-PLS on eighty-five calibration seeds. Error bars indicate 1 × standard deviation of multiple models made on subsets of recordings with the indicated seed orientations.

spectra with second derivative + MSC [30,31]. This combination showed a prediction error of 0.48 %dm protein using 5 PLS components, while EISC pre-treated model had a similar prediction error of 0.49 %dm protein at the expense of being less parsimonious, however, using 7 PLS components. That particular investigation was based on a calibration set of 415 wheat seeds and a test set of 108 seeds and further used in developing a screening method for characterization of single-kernel wheat [14]. None of these studies, however, assessed sample presentation as a possible source of error.

Because a single seed is itself a heterogeneous unit with respect to its internal constitution, measuring it non-destructively necessitates that the near-infrared light should transmit through as much of the seed interior as possible. If this cannot be achieved by a single recording, multiple scans from different orientations would presumably work towards better representativity of the ultimate averaged NIT spectrum. This was indeed similar to the hypothesis of Delwiche [16] who suggested that it might be possible to improve model error by averaging NIT recordings of the seed presented in various orientations.

In Fig. 3 the percentage of protein RMSECV is presented as a function of the different sample presentation strategies realised. The first eight strategies, which would be among the fastest, are when a single orientation is used for all recordings. Error bars indicate the standard deviation of the RMSECV derived by running the model several times on subsets of the orientation in question. Two orientations have RMSECV significantly lower than the others: the furrow down and the germ pointing towards the centre or the edge of the cassette respectively. Averaging over the three replicates of each orientation improved markedly on most of the models, but still showed the orientations with the furrow down as the best (bar No. 10 and 14 indicated with black and grey bars in the figure) — each validating the other, since pointing the germ towards the centre or the edge is a symmetrical change not expected to alter the spectra.

The last five strategies are from left to right as indicated by the seeds in Fig. 3: A single recording for each seed in a random orientation; 2: averaging two random orientations subject to the constraint that one orientation is with the furrow either to the left or to the right and the other with the furrow up or down; 3: averaging four orientations disregarding the direction of the germ; 4: averaging over all eight orientations; 5: average of all 24 recordings of every single seed. The last one indicated by the black bar is the only one to compete with strategy No. 10, both with an RMSECV of only 0.36, far less than has previously been found [14–16] where sample presentation has not been addressed in detail. Why one orientation is better and faster than averaging all orientations is not clear and should investigated further and apparently rejects the hypothesis of Delwiche [16]. However, the origin of this paradox may be found in the surface characteristics of the individual seeds. The opposing site to the furrow is the smoothest and most uniform surface of wheat seeds in general. When the near-infrared light meets the surface one may hypothesise that a more uniform scatter pattern emerge which in turn increases the effectiveness of the scatter correction employed. Light scatter from more fractured surfaces — especially the furrow side, may be more difficult to handle, thus leading to larger model error when used either alone or in averages with other orientations.

The model using four PLS components associated with optimal seed orientation, also shown in Fig. 2, was used for recording all the 1589 remaining seeds (including the four reference seeds) in triplicate in order to predict their protein content and heterogeneity. Considering the relatively small model based on only 85 random seeds, remarkably few seeds were discarded as outliers, only 86 of 1680, 33 of which were from instrumental errors, 19 from operational errors and the rest due to poor model fit. Discarded outliers were not replaced, since the seeds were selected completely by random in quite different lots and not by their appearance and was as such an expected outcome of the sampling. Obviously not all of the variability found in the lots would be accounted for in the prediction model, thus leading to a few unpredictable outliers to be discarded.

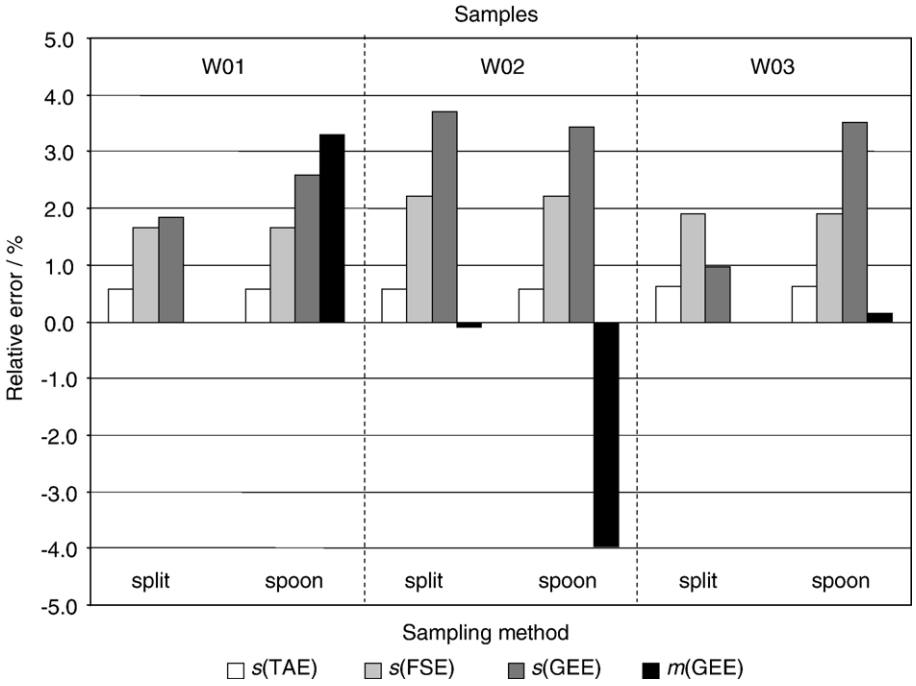


Fig. 4. Comparison of the relative errors from estimating the protein content of three wheat lots from forty-two seeds selected by splitting or by spoon sampling. The relative errors of TAE and FSE are from Table 1 and are presented here again for comparison; they are identical for both split and spoon sampling. Bias is calculated by assuming the concentrations from Table 1 are true concentrations, following [11].

For the prediction error RMSECV was used, though rounded up to 0.4 %dm protein, keeping in mind that no test set was made and that LOO cross validation *per force* gives rise to overoptimistic prediction errors [26]. However, prediction errors using test sets in other robust single seed investigations [14–16] were not markedly different from the cross validation errors. And even though test sets are preferred, they are not crucial for this investigation, since the analytical error is an order magnitude lower than the sampling errors encountered below and a potentially over- or underestimation will have no severe effect on the interpretations and conclusions made.

A summary of the results from the thirteen lots mass-reduced to forty-two seeds by splitting is presented in Table 1. Since the sampling of W01, W02 and W03 was repeated five times, 210

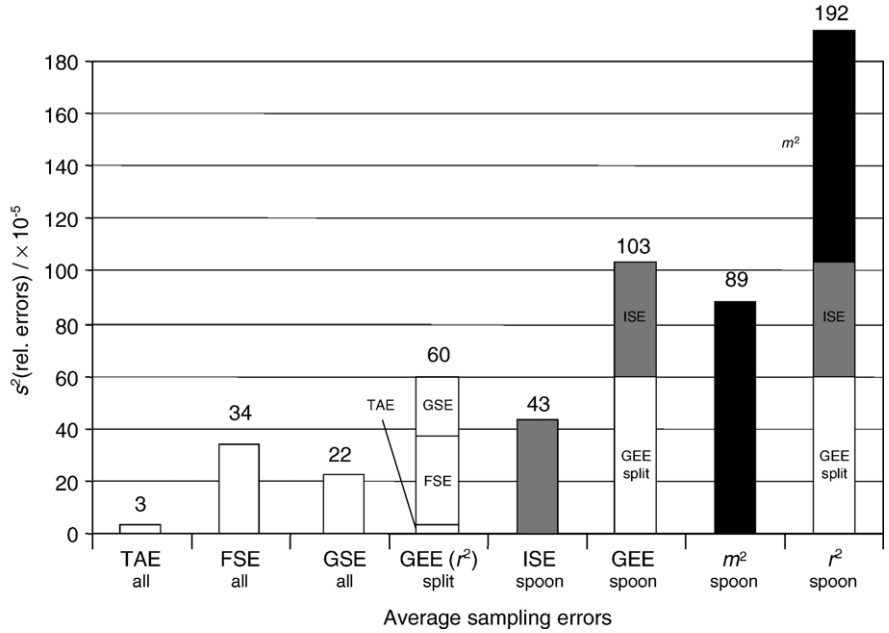


Fig. 5. Breakdown of the average GEE variance for samples of W01, W02 and W03 for both spoon and split sampling. The squared bias and the representativity of spoon sampling are also presented, the latter to be compared to the representativity of the split sampling in that s^2(GEE split)=r^2(split).

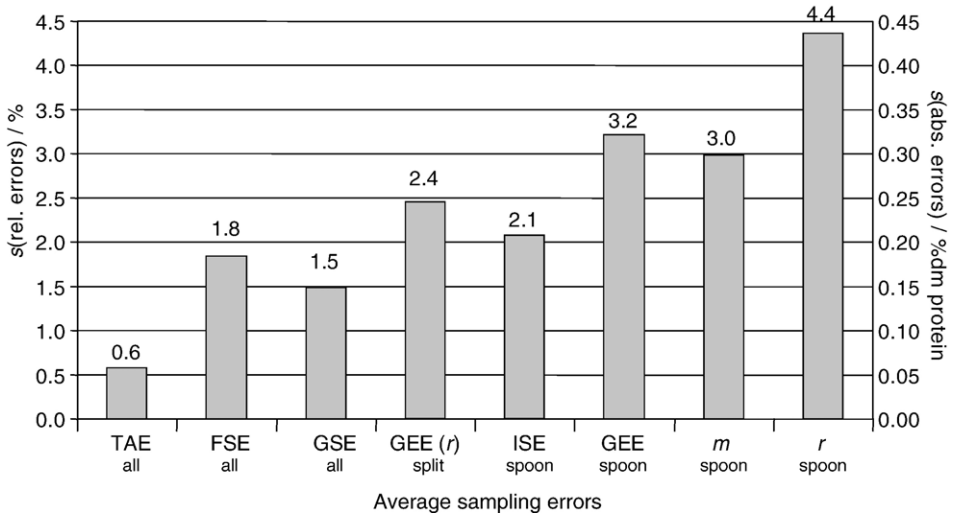


Fig. 6. The same errors as presented in Fig. 5 in relative and absolute values, the latter assuming a protein concentration of 10 %dm to be related to the units of the analyte.

seeds were used for calculations here, while only 42 seeds were used for the remaining samples. Disregarding the outliers, the following parameters are presented; M_{i*}, the bulk protein content, a_L, the single seed variances, s^2(FSE) and standard deviations, s (FSE) (relative) and s (absolute %dm protein), constitutional heterogeneity (Eqs. (13) and (14)), CH_L and the derived fundamental errors of a forty-two seed size sample.

It is noteworthy that the heterogeneity of the individual samples varies markedly. The masses and single seeds ranged from 6.2 mg to 79.0 mg with a dry matter protein content varying from 6.9%–22.5%. As shown by Delwiche [17], there is no useful relation between kernel mass and protein content — however, it cannot be stated that the protein content is randomly distributed along the different masses. Bulk protein concentrations vary from 9.9 to 15.4 %dm protein with average seed masses ranging from 33.9 to 52.9 mg. The standard deviation of the single seed protein varies from 11 to 21% relative and 1.1 to 2.5 %dm protein. The analytical error (RMSECV = 0.4) from the PLS-R model is indicated in the last line and, as expected, constitutes only a fraction of the seed variation. The constitutional heterogeneity used to calculate the fundamental error of samples with a higher number of seeds also varies correspondingly from as low as 8.8×10^{-3} (W04) to as high as 46.6×10^{-3} (S10). The fundamental standard deviation of a 42 seeds sizes sample thus varies from 0.2 to 0.4 %dm protein.

The relative standard deviation and biases of the first three samples are presented in Fig. 4 for both split and spoon (grab) sampling strategies. For all strategies the standard deviation of the global estimation error from five replicates are shown (dark grey) together with the fundamental sampling error (light grey) as well as the total analytical error (white) from Table 1. Since FSE is limiting GEE, the standard deviations found should all be higher than s(FSE) [1,3]. This is also the case in all but one case, W03 split, which probably is due to the very few samples extracted — only five. Although we might expect the grab-sampled (spoon) samples to be much higher in s(GEE), this cannot clearly be deducted from this figure. Even though W01 and W03 demonstrate such behaviour, W02 does not. However, the spoon samplings of W01 and W02 are remarkably biased, while this is not the case for the rest. It must be noted that bias is calculated on the necessary assumption that results from the splitter are unbiased and the true protein concentration is the one found in Table 1; however, this assumption would appear to be amply substantiated [22]. Strictly speaking, bias should then be understood as difference between spoon- and split-sampled mean values. Overall the figure shows that spoon samples have either high standard deviation or high bias or both compared to split samples, which is expected according to TOS [1–6,8,11,22].

In order to break down the variance into its additive contributions, the squared errors in Fig. 4 were averaged and subsequently presented in Fig. 5. It is clearly seen that $s^2(GEE_{split})$ is the sum of three contributions $s^2(TAE) = 3 \times 10^{-5}$, $s^2(FSE) = 34 \times 10^{-5}$ and $s^2(GSE) = 22 \times 10^{-5}$, the latter found by subtraction using Eq. (15). $s^2(GEE_{spoon})$ has an extra contribution from $s^2(ISE) = 43 \times 10^{-5}$ also found by subtraction (Eq. (16)). Clearly all three sampling errors have a significant impact on the global estimation error. They are all of the same order of magnitude, which is approximately ten times the total analytical error (TAE).

In addition, the squared bias, $m^2 = 89 \times 10^{-5}$, is introduced in the error balance, to be added to the $s^2(GEE_{spoon})$ to form the representativity score (here called a *score*, since it emerges from two variables, the standard error and the bias, that are usually not added together, but can be used to rank different techniques or strategies [22]). Since a lower representativity score is optimal it is very easy to see that the unbiased splitting strategy (as per representative sampling) for this material ($s^2(GEE_{split}) = r^2_{split}$) has a significantly lower magnitude compared to that of the spoon sampling. Clearly spoon sampling or grab sampling both increase sampling variance and introduce bias of unknown size. This is in accordance with other studies on sampling strategies of other particulate material [22–24] and thus stresses the importance of representative sampling even in apparently homogeneous lots.

Others have suggested the use of *appropriate sampling* rather than *representative sampling* as a better way to assess estimation

uncertainties [10]. This distinction between appropriate and representative is rather artificial, however, since the causes of variance and biases in analytical results are agreed upon and representative sampling must be the appropriate sampling strategy whenever it is economically sound and indeed physically possible and desirable, i.e. whenever the lot is or can be turned into a zero- or one-dimensional lot [11].

Since variances are difficult to interpret quantitatively, apart from their additive properties, the result from Fig. 5 is presented in Fig. 6 as standard deviation both on a relative scale and on the absolute scale, allowing assessment in the original measurement units. From Table 1 the impact of FSE can easily be enlarged with standard deviation of up to 4%. This in turn will have an impact on GSE which is related to FSE and thus will also impact GEE significantly, even though sampling is conducted in a representative way. As suggested by Delwiche and Hruschka [18], near-infrared reflectance spectroscopy of single seeds may be a good way to characterize the protein concentration of bulk samples. They suggested using averaged spectra from 300 seeds (canonical number) in order to determine the bulk protein concentration with a standard deviation of 0.25% protein (12% moisture basis). Apparently our studies suggest that 42 seeds would actually do the job just fine, even incorporating the significant sampling errors (Fig. 6) which were acknowledged as being one of the large unknowns in Delwiche and Hruschka's study [18].

However, the fundamental error in their study was larger — comparable to W09 and S10 (Table 1), and they did not include the grouping and segregation error which is shown here to be of comparable magnitude to the fundamental error. They instead left the model error playing a greater role, since they averaged the NIR spectra and protein contents before modelling. In our study the RMSECV is a part of the overall kernel variance and is as such diminished the more seeds are accumulated in the sample. Nevertheless, theirs and our studies both point to using approximately 300 seeds, when the fundamental error (approximately the standard deviation between seed protein concentration) can be as high as 3 %dm protein and the grouping and segregation error is estimated to be of the same magnitude. With sampling errors of this order there is no reason to perform scans in triplicate. To generalize from the present work, and that of Delwiche, we tentatively suggest a constitutional heterogeneity of protein in wheat to be expected up to the order of 50×10^{-3}.

Usually CH_L is not experimentally available in particulate material, in which case DH_L can be used to assess an appropriate sample size as well as the number of increments needed for a representative sample. Rough estimates of the distributional heterogeneities, i.e. the weighted variances of 42 seed-sized samples are presented in Table 2. The average value of 64×10^{-5} is, however, not a result to rely upon in future studies, since it inherently will vary a lot from lot to lot depending on both the size of the fundamental error as well as the grouping and segregation in the actual investigation.

Table 2
DH_L estimated based on five split sampled increments of 42 seeds of samples W01, W02 and W03

ID	# Inc.	M_n* [mg]	DH_L [$\times 10^{-5}$]
W01	5	1950.0	37
W02	5	1620.0	145
W03	5	1630.0	9
		Average:	64

5. Conclusions

From studies of two alternative sampling strategies, representative splitting *versus* direct grab sampling, in thirteen different wheat lots it is concluded, with emphasis, that representative sampling is a must, even in apparently homogeneous materials like ordinary wheat lots.

Quantification of all additive and inevitable sampling errors shows that both the fundamental (FSE) and the grouping and segregation error (GSE) have a significant impact compared to analytical errors in the present material. Both the variance of FSE and GSE were an order of magnitude larger than the variance of TAE. The variance of the incorrect sampling errors (ISE) and the squared *bias* from grab sampling were found to be even larger, thus inflating the variance of GEE and the representativity score alarmingly. Even for the exceedingly uniform grain-type material, heterogeneities are highly significant — and thorough sampling considerations deal effectively in reducing or eliminating all these sampling errors.

Moreover indications were observed that presenting single seeds in a certain way, i.e. with the elongated furrow pointing towards the detector in the NIT beam, leads to prediction models with significantly lower prediction error compared to all other strategies for presenting the seeds. This thought-provoking result is hypothesised to be due to the different scatter patterns induced by the surface structure of the seeds.

Acknowledgements

We thank The Royal Veterinary and Agricultural University for the first author's Ph.D. scholarship and thus the funding of this work. We are grateful to Lisbeth Hansen who thoroughly performed all reference measurements. Special thanks to Q-Interline for the Sampling Award 2005 acknowledging the TOS work presented here.

References

[1] P. Gy, Sampling for Analytical Purposes, John Wiley and Sons, 1998.
[2] P.M. Gy, Trac-Trends Anal. Chem. 14 (1995) 67–76.
[3] P. Gy, Chemom. Intell. Lab. Syst. 74 (2004) 7–24.
[4] P. Gy, Chemom. Intell. Lab. Syst. 74 (2004) 25–38.
[5] P. Gy, Chemom. Intell. Lab. Syst. 74 (2004) 39–47.
[6] P. Gy, Chemom. Intell. Lab. Syst. 74 (2004) 49–60.
[7] F. Lamé, T. Honders, G. Derksen, M. Gadella, Environ. Pollut. 134 (2005) 5–11.
[8] K.H. Esbensen, K. Heydorn, Chemom. Intell. Lab. Syst. 74 (2004) 115–120.
[9] K. Heydorn, E.H. Hansen, Chemom. Intell. Lab. Syst. 79 (2005) 129.
[10] M.H. Ramsey, Accredit. Qual. Assur. 7 (2002) 274–280.
[11] L. Petersen, P. Minkkinen, K.H. Esbensen, Chemom. Intell. Lab. Syst. 77 (2005) 261–277.
[12] H.P. Song, S.R. Delwiche, Y.R. Chen, Opt. Eng. 34 (1995) 2927–2934.
[13] S.R. Delwiche, D.R. Massie, Cereal Chem. 73 (1996) 399–405.
[14] J.P. Nielsen, D.K. Pedersen, L. Munck, Cereal Chem. 80 (2003) 274–280.

[15] D.K. Pedersen, H. Martens, J.P. Nielsen, S.B. Engelsen, Appl. Spectrosc. 56 (2002) 1206–1214.
[16] S.R. Delwiche, Cereal Chem. 72 (1995) 11–16.
[17] S.R. Delwiche, J. Cereal Sci. 27 (1998) 241–254.
[18] S.R. Delwiche, W.R. Hruschka, Cereal Chem. 77 (2000) 86–88.
[19] S.R. Delwiche, F.E. Dowell, Getreide Mehl Brot 56 (2002) 141–146.
[20] S.R. Delwiche, Trans. ASAE 36 (1993) 1431–1437.
[21] E.B. Maghirang, F.E. Dowell, Cereal Chem. 80 (2003) 316–322.
[22] L. Petersen, C.K. Dahl, K.H. Esbensen, Chemom. Intell. Lab. Syst. 74 (2004) 95–114.
[23] R.W. Gerlach, D.E. Dobb, G.A. Raab, J.M. Nocerino, J. Chemom. 16 (2002) 321–328.
[24] P.L. Smith, Chemom. Intell. Lab. Syst. 74 (2004) 225–230.
[25] D. Adams, The Hitch Hiker's Guide to the Galaxy. A Triology in Five Parts, William Heinemann, London, 1995.
[26] K.H. Esbensen, Multivariate Data Analysis — In Practise. An Introduction to Multivariate Data Analysis and Experimental Design, 5th Ed.CAMO AS Publ., Oslo, Norway, 2001, 600 pp.
[27] T. Næs, T. Isaksson, T. Fearn, T. Davies, A User-Friendly Guide to Multivariate Calibration and Classification, NIR Publications, Chichester, UK, 2002.
[28] L. Nørgaard, A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, S.B. Engelsen, Appl. Spectrosc. 54 (2000) 413–419.
[29] H. Martens, J.P. Nielsen, S.B. Engelsen, Anal. Chem. 75 (2003) 394–404.
[30] P. Geladi, D. Macdougall, H. Martens, Appl. Spectrosc. 39 (1985) 491–500.
[31] O.E. Denoord, Chemom. Intell. Lab. Syst. 23 (1994) 65–70.

Bulk quality diversification of organic wheat by single-kernel near-infrared (SKNIR) sorting

E. Tønning [ab,*], A. K. Thybo [b], L. Pedersen [c], L. Munck [a], Å. Hansen [a], S. B. Engelsen [a], L. Nørgaard [a]

[a] *Quality & Technology, Department of Food Science, Faculty of Life Sciences, University of Copenhagen, DK-1958 Frederiksberg C, Denmark*[1]

[b] *Plant Food Science Group, Department of Food Science, Faculty of Agricultural Sciences, University of Aarhus, DK-5792 Aarslev, Denmark*[2]

[c] *Department of Chemical Engineering, Faculty of Engineering, University of Southern Denmark, DK-5230 Odense, Denmark*[3]

* Corresponding author:

Erik Tønning

Plant Food Science Group, Department of Food Science, Faculty of Agricultural Sciences, University of Aarhus, DK-5792 Aarslev, Denmark

Tel.: +45-8999 3413; Fax: +45-8999 3495; *E-mail address*: erik.tonning@agrsci.dk

[1] http://www.models.life.ku.dk

[2] http://www.agrsci.org

[3] http://www.sdu.dk

Abstract

This paper explores the effects of fractioning heterogeneous bulk wheat by fast single-kernel near-infrared (SKNIR) sorting according to an internal complex quality trait using a fast prototype object sorter. The effect of sorting was compared to diversification by means of varying the preceding catch crops in an organic field experiment in two growth years and two locations. In the resulting twenty-nine wheat lots, fifty-two quality parameters were measured on grains, flours and breads. The data was analysed by principal component analysis (PCA) and analysis of variance (ANOVA). Within each year and location/cultivar the SKNIR fractionation had a significant effect on bulk grain density, protein and wet gluten content, Zeleny sedimentation volume, Farinograph water absorption, Farinograph softening, falling number and gelatinisation temperature. In comparison, varying the preceding catch crops did not show any coherent systematic. This investigation shows that the fast SKNIR sorter can qualify a given heterogeneous wheat grain lot and potentially improve the market and functional value post harvest especially with emerging sorting equipment capable of sorting several tonnes per hour.

Keywords: Wheat quality, single-kernelnear-infrared sorting, quality fractionation, heterogeneity, functionality, baking, breeding, SKNIR, NIT, chemometrics, grain, flour, dough, bread, organic farming, catch crops

List of abbreviations used: ANOVA = analysis of variance; A-PLS = ANOVA-PLS; DON = deoxynivalenol; NIT = near-infrared transmission; PC = principal component; PCA = principal component analysis; PLS = partial least squares; SKNIR = single-kernel near-infrared; TPA = texture profile analysis

Introduction

The quality of wheat is presently controlled by factors directly related to the primary production conditions such as choice of cultivar, soil quality, weather, availability and quality of fertilisers, pressure from pests, i.e. weeds, insects and diseases. In addition organic farmer restrictions regarding mineral fertilizers and the use of pesticides significantly reduce the possibilities of ensuring a specific crop quality at the time of harvest (Frederiksson et al., 1997; Johansson et al., 2001; Triboi and Triboi-Blondel, 2001). Apart from the consumer preference for products produced with minimal environmental impact the organic crop may thus be suited for anything from energy source to industrial bread production depending on internal chemical and physical quality and uniformity (Haglund et al., 1998; Kihlberg et al., 2004).

More attention has recently been drawn to the fact that even though a batch of plant material, such as wheat, appears uniform there exist significant variation in the quality of individual grains, e.g. the protein content may vary from 5 to 20% in some wheat lots (Delwiche and Hruschka, 2000; Pedersen et al., 2002; Tønning et al., 2006). If sorted according to internal quality a potential diversification may increase the value and uniformity of wheat from both organic and conventional farming. Single-kernel near-infrared (SKNIR) spectroscopy in combination with chemometric tools such as partial least squares (PLS) has been utilised to predict protein, vitreousness, density and hardness of single kernels of wheat (Delwiche, 1993, 1995, 1998; Nielsen et al., 2003, Tønning et al., 2006) and may serve as a relevant technology for such endeavours. Baking quality and dough handling properties can be predicted in bulk ground wheat (Delwiche et al., 1998; Delwiche and Weaver, 1994) and is already determined routinely from near-infrared transmission (NIT) spectra recorded in commercially available instrumentation such as the Foss Infratec™

1241 Grain Analyser with built-in calibrations for parameters such as Farinograph water absorption, stability, wet gluten and Zeleny sedimentation volume in addition to the standard protein and moisture calibrations. Developments in SKNIR systems for fast determination of attributes and sorting are now in the making. Dowell et al. (2006) reported a system which sorted wheat kernels with respect to protein and hardness respectively and prose millet according to amylase content. Sorting wheat with respect to infestation with Fusarium graminearum which pose a health risk due to production of the mycotoxin deoxynivalenol (DON) was successfully conducted by Delwiche et al. (2005) using high speed optical sorting. Pasikatan and Dowel (2004) tried to separate mixtures of high and low protein wheat kernels on a commercial colour sorter equipped with near-infrared filters. Rittiron et al. (2004) made an initial suggestion for a brown-rice sorting machine with respect to moisture and dry matter protein content. Also for corn and oil-seed a SKNIR system for fast determination of quality attributes has been developed (Armstrong, 2006). Fractionation by other means such as physical appearance, size and weight is an alternative to the internal quality trait determined by near-infrared technology (Elfverson et al. 1999; Yoon et al., 2002).

A new industrial "TriQ" SKNIR sorting system with a capacity of sorting several tonnes per hour utilising the heterogeneity of bulk wheat is coming to the market (Löfqvist and Nielsen, 2003; Löfqvist and Nielsen, 2006). It diversifies bulk crops post harvest by sorting the material kernel by kernel using SKNIR technology calibrated to internal quality. The ultra fast TriQ sorting system arranges and fix each individual grain into separate perforated depressions in a rotating drum in such a way that the quality of each individual grain can be discretely determined and subsequently be discretely handled according to its quality. Preliminary results from a prototype sorter has been

presented in which a wheat lot with an average dry matter protein content of 12.3% was sorted according to protein content <11%, 11-13% and >13% (Nielsen, 2002). The relative yields were 26%, 38% and 36% with average protein contents of 10.2, 12.0 and 14.4%, respectively. Increased protein content and quality is well known to improve loaf volume (Bushuk et al., 1969).

Post harvest catch crops may immobilise, preserve and accumulate nutrients in the top soil in order to prevent unnecessary loss of valuable nutrients. Catch crops have been shown to prevent leaching of nutrients – especially nitrogen as nitrate or ammonia dissolved in the soil water (Eriksen et al., 2006; Francis et al., 1992). Leaching nutrients may potentially harm the recipient such as naturally nutrient deficient ecosystems or water supply areas. Ideal catch crops are able to cover the post harvest field very fast and utilise available nutrients by developing an extensive and deep root system in order to catch leaching nutrients and bring them to the surface before they are completely lost to the recipient (Kristensen and Thorup-Kristensen, 2004). Especially in organic farming where nutrients are deficient, expensive and less available for the plants the use of legumes and post harvest catch crops with deep root systems may in addition to the environmental effect also affect the quality of the subsequent crop (Eriksen et al., 2006).

The aim of this investigation was to explore the quality diversification effects of fast post harvest SKNIR sorting of organic wheat compared to the effects of varying the preceding catch crops. Two different catch crops, winter vetch and fodder radish or turnip were used for differentiating organically grown wheat quality as compared to no catch crop in two growth years and two locations. On top of this, a laboratory scale prototype TriQ SKNIR sorting system was tested on the same material by which each grain was sorted into three different quality

fractions (Löfqvist and Nielsen, 2003). The quality of the field lots and SKNIR fractions were determined in a multi-step investigation of resulting grains, flours and breads. Fifty-two quality parameters were measured ranging from protein content in the grain lots over instrumental texture of the bread to sensory panel evaluations, i.e. from grain to brain.

Experimental

Materials

Twelve lots of organically grown spring wheat from factorial Danish/German field trials in 2003 (03) and 2004 (04) were used and further diversified by SKNIR sorting. Growth sites were Aarslev (Aa) in Denmark and Kiel (Ki) in Germany, respectively. Two spring cultivars, Vinjett (Vi) and Combi (Co) were used. Three different catch crops were grown prior to cultivation of the spring cultivars. The catch crops were: Winter Vetch (*Vicia villosa*) (WV), Fodder Radish (*Raphanus sativus var. Oleiformis*) (FR) or Turnip (*Brassica rapa*) (Tu). The last two are both from the *Brassicaceae* family and were used interchangeably. Combi was available from Kiel, 2003 and 2004 using Winter Vetch, Fodder Radish or no catch crop (NC). Vinjett was available from Aarslev, 2003 and 2004 using Winter Vetch, Fodder Radish/Turnip or no catch crop. Each of the fourteen lots was accumulated from three plots and varied in size from 7.9 kg to 18.2 kg subject to availability (Table 1). Climate, soil properties and yield were not subjects of interest in this investigation and will not be reported here.

From each growth year and each location approximately half of the grains from each catch crop treatment were mixed forming a mixture treatment (Mx) (Table 1). The mixtures were then sorted grain by grain into three approximately equally sized fractions representing a low (F1), medium (F2) and high (F3)

baking quality using laboratory scale TriQ SKNIR sorter (BoMill AB, Lund, Sweden) with a capacity of 2 kg h^{-1} (Table 1). The TriQ sorter recorded a near-infrared reflection spectrum of each kernel from 1100 nm to 1700 nm with a bandwidth of 10 nm using a tungsten lamp and a diode array spectrometer. A calibration PCA model based on raw spectra from thousands of kernels of various origins had previously been established (Löfqvist and Nielsen, 2003). The entire lot of grains was subsequently sorted by classification based on a combination of the scores on the first two latent variables calculated from the spectra. The complex quality targets were set individually for each mixture in order to achieve approximately equally sized fractions (F1, F2 and F3) (Table 1).

The mixing and subsequent fractionation resulted in four mixture lots and twelve (4 × 3) sorted lots based on the distribution of qualities within each mixture. The mixture lot from Kiel, 2003 however was excluded in the subsequent process due to a very small lot size (0.5 kg). Two additional winter cultivars, Capo (Capo) available from Kiel, 2004 only and Pentium (Pent) available from Aarslev, 2004 only, both grown without preceding catch crop was included for reference. In total twenty-nine lots varying in size from 3.7 to 15.2 kg secured enough material for the entire processing of grain to flour and bread with subsequent measurements (Table 1). The abbreviations in brackets above were used to form lot names for easy identification, thus a Vinjett lot from Aarslev, 2004 grown after Fodder Radish has ID: 04AaViFR.

Table 1: Experimental design with two locations/cultivars, AaVi and KiCo and two growth years, 03 and 04. In each block three different preceding catch crop strategies, NC, FR/Tu and WV were used. A mixture (Mx) of these was subsequently fractionated by the TriQ system, F1, F2 and F3 (Arrows). The mixture proportions as well as yields in the approximately equally sized fractions are indicated. Two winter cultivars with no preceding catch crop were also included. Flour yield and dry matter ash content of the flour is shown for reference.

Lot	M_{mix} [kg]	Sorting	M_{res} [kg]	Fl. Yield [%]	Ash DM [%]
2003 Aarslev Vinjett (03AaVi)					
No Catch (NC)	3.6 (20.8%)		4.1	71.3	0.53
Fodder Radish (FR)	6.0 (34.7%)		6.4	68.2	0.51
Winter vetch (WV)	7.7 (44.5%)		8.2	67.6	0.50
Mixture (Mx)	17.3 (100%)	→	3.8	71.8	0.61
Fraction 1 (F1)			4.1 (30.4%)	68.3	0.49
Fraction 2 (F2)			4.5 (33.3%)	68.1	0.50
Fraction 3 (F3)			4.9 (36.3%)	68.2	0.52
2004 Aarslev Vinjett (04AaVi)					
No Catch (NC)	6.7 (29.3%)		7.1	65.9	0.51
Turnips (Tu)	8.9 (38.9%)		9.4	66.6	0.51
Winter vetch (WV)	7.3 (31.9%)		7.8	69.2	0.53
Mixture (Mx)	22.9 (100%)	→	5.4	66.8	0.58
Fraction 1 (F1)			5.9 (35.1%)	69.0	0.50
Fraction 2 (F2)			5.5 (32.8%)	68.9	0.50
Fraction 3 (F3)			5.4 (32.1%)	69.4	0.51
2003 Kiel Combi (03KiCo)					
No Catch (NC)	5.3 (32.7%)		5.6	68.9	0.50
Fodder Radish (FR)	5.4 (33.3%)		5.9	68.1	0.49
Winter vetch (WV)	5.5 (34.0%)		5.8	66.1	0.49
Mixture (Mx)	16.2 (100%)	→	0.5 *	n.a.	n.a.
Fraction 1 (F1)			4.9 (31.6%)	68.8	0.50
Fraction 2 (F2)			5.5 (35.5%)	69.3	0.51
Fraction 3 (F3)			5.1 (32.9%)	71.9	0.52
2004 Kiel Combi (04KiCo)					
No Catch (NC)	7.1 (31.0%)		7.1	66.5	0.48
Fodder Radish (FR)	8.4 (36.7%)		8.3	66.1	0.49
Winter vetch (WV)	7.4 (32.3%)		7.4	66.3	0.49
Mixture (Mx)	22.9 (100%)	→	4.7	68.6	0.52
Fraction 1 (F1)			5.3 (30.8%)	72.3	0.51
Fraction 2 (F2)			6.3 (36.6%)	70.3	0.49
Fraction 3 (F3)			5.6 (32.6%)	69.3	0.51
2004 Winter Crops					
04AaPent			16.2	65.4	0.57
04KiCapo			11.2	67.3	0.47

* This lot was too small and not included in further analysis

Milling

The lots were milled on a laboratory scale mill, Brabander Quadromat Senior (Duisburg, Germany) separating the bran and germ from the endosperm. The process involved a conditioning step in which 0.7 to 2.2% of water was added depending on the actual moisture content of the grains. Grains and water were mixed in a rotating drum for 2.5 to 8.5 hours in order to soften the bran prior to the milling. The flour yield varied from 65.4% to 72.3% with 99.3% to 99.8% of the flour in particle sizes smaller the 160 µm (sieve size). Flour yields and ash contents are reported in Table 1 for reference. The flour lots were kept in airtight plastic containers.

Characterisations

Physicochemical characterisations throughout the entire process from grain lots to bread were carried out in the following steps. A list of all parameters measured with a brief description can be found in Table 2.

Grain NIT analysis

Approximately 800 g of grain, representative for the lot, was poured into a Foss Infratec™ 1241 Grain Analyser. Ten NIT spectra from 850 - 1048 nm for every 2 nm were recorded and averaged. Dry matter protein, bulk grain density, moisture content, Zeleny sedimentation volume, dry matter starch content and wet gluten content was automatically calculated based on the built-in global Foss calibration.

Flour NIT analysis

Moisture, dry matter protein, dry matter ash, Farinograph water absorption, Farinograph stability and wet gluten content were predicted by Foss Infratec™ 1241 Grain Analyser using the Foss world-wide calibration. Two Infratec™

flour cups were filled for each lot with flour and inserted into the hopper. For each cup ten near-infrared transmission scans from 850 nm to 1048 nm were recorded and averaged. Results from the replicate measurements were averaged.

Standard flour test

The falling number was determined in duplicates according to ICC Standard No. 107/1 and averaged. Farinograph water absorption, development time, stability and softening were determined in a 50 g mixing chamber according to ICC Standard No. 115/1. Zeleny sedimentation volume was determined in duplicates according to AACC Standard No. 56-61A and ICC Standard No. 116/1 and averaged. Amylograph gelatinisation maximum and temperature at maximum gelatinisation was determined in duplicates according to ICC Standard no. 126/1 and AACC Standard No. 22-10 and averaged. Wet gluten and gluten index were determined in duplicates according to AACC Standard No. 38-12 and ICC Standard No. 155 and averaged.

Dough preparations and bread baking

For each lot, three breads were made in household baking machines (Dubuc and Boudreau, 1992; Grausgruber et al., 2001; Hansen and Hansen, 1992, 1993; Peltonen and Salovaara, 1991; Zwingelberg and Brümmer, 1990), Panasonic Automatic Bread Bakery, SD-253 in the following manner. 7.1 g of dry yeast was distributed evenly in the pan around the kneading blade. Then 600 g of flour on a 14% moisture base were added followed by 8.4 g table sugar and 8.4 g salt. 5 mL 0.48% ascorbic acid solution corresponding to 40 mg ascorbic acid/kg flour and tap water at 25ºC up to a total water amount corresponding to Farinograph water absorption (500 FU) was added. The baking process was initiated using the following SD-253 program settings, Size: XL, Bread type selection: BASIC, Option: BAKE RAPID and Crust colour: MEDIUM. The

basic program ran for 115 min. as follows: 20 min. kneading, 20 min. rising, 1 min. kneading, 39 min. rising and 35 min. of baking.

Immediately after baking, the breads were removed from the baking pans and left on a wire rack to cool for 90 minutes. The breads were weighed and the longitudinal as well as the transverse circumference, c_L and c_T, over the bread centre were measured. The volume of the bread was estimated by assuming a cuboid shaped bread with length and width determined by the pan size, 175 mm × 130 mm, and the height determined from circumferences:

$$\text{Height} = \frac{c_L / 2 - 175\text{mm} + c_T / 2 - 130\text{mm}}{2} \qquad \text{(Eq. 1)}$$

This cuboid assumption does not take the individual more or less rounded shapes with cracks and pores of the bread crusts into account, but serves as an adequate substitute for the actual volume and for density calculation. Additionally the relative water loss during baking and cooling was registered. Two breads were frozen to -18ºC for texture and sensory analysis of thawed bread, while one was used immediately for texture analysis of the fresh bread crumb.

Texture profile analysis of fresh bread crumb

From each lot, one bread was cut into nine approximately 20 mm slices on a slicing machine. Actual slice thicknesses varied between 17.9 and 23 mm. Only slice no. 2, 3, 7 and 8 were used for texture profile analysis (TPA). The end slices 1 and 9 as well as centre slices 4, 5 and 6 were discarded to avoid influences from the crust and the kneading blade perforation respectively. The TPA was conducted by placing an entire slice of bread on the base plate of the TA-XT2 Texture Analyser, Stable Micro Systems fitted with a cylindrical 40

mm SMS steel probe with sharp edges. The force used to compress and decompress the bread crumb 15 mm (75%) twice with a 5 s delay between the first and the second compression was recorded. Test speed was 1 mm s^{-1}, trigger force 0.02 N and post test speed was 5 mm s^{-1} after second down stroke. From each time-force curve and actual slice thickness the following nine parameters were extracted and calculated (Table 2) and the results from the four slices were averaged:

Hardness 1 at 60% compression (H1): Force at 60 % compression during first down stroke corrected for deviation from 20 mm slice thickness. H1 = $F_{60\%,20\text{mm}}$. The force at 60% compression, $F_{60\%}$, and thickness, d in mm, were given by the TA-XT2. Thus the force expected to be used on the individual slices if they would have been 20 mm, $F_{60\%,20\text{mm}}$, can be calculated:

$$F_{60\%,20\text{mm}} = F_{60\%} \frac{20\text{mm}}{d} \qquad \text{(Eq. 2)}$$

Hardness 2 at 60% compression (H2): Determined in the same way as H1 above, only for second down stroke. The 60% compression corrected with sample thickness was used as a more robust hardness measures in stead of the less robust maximum compression (75%) which due to the high compression rate is very sensitive to deviations in sample height. Adhesive Force (AF): The maximum negative force exhibited during first up stroke. Adhesiveness (Ad): The work done by the bread crumb to hold on to load cell, i.e. the negative area under the curve, during first up stroke. Resilience (Re): Area under curve during first upstroke until zero force divided by area under curve during first down stroke. Cohesiveness (Co): Area under curve during second down stroke divided by area under curve during first down stroke. Springiness (Sp): Distance travelled during second down stroke divided by distance travelled under first

compression (15 mm). Gumminess (Gu): Gu = Co·H1. Chewiness (Ch): Hh = Gu·Sp.

Texture profile analysis and sensory evaluation of thawed bread crumb

One bread was thawed at room temperature and equilibrated for 16 hours before cut into 20 mm slices. One slice was used for TPA on a Stevens Metric QTS 25 connected to a Windows personal computer running Stevens Farnell QTS25 Profile version 1.1WT 14-May-96. The test speed was 1 mm/s and trigger point 0.05 N using the same probe as for fresh bread. The parameters hardness 1 (H1) and hardness 2 (H2) were the maximum force used at 75% compression (15 mm) during first and second compression as no information on actual sample thickness was available from the QTS25. The rest of the parameters were extracted and calculated as for fresh bread (Table 2). The remaining slices were used for sensory profiling. Sensory profiling was performed in a sensory evaluation laboratory according to international standards (ASTM STP 913). The basic trained panel with many years of sensory experience was composed of 10 assessors (4 males/6 females, aged from 30-55 years). They were trained for 6 hours in sensory profiling of bread. Bread slices were cut vertically in halves and served to the assessors. The twenty-nine bread samples were served in 5 sensory sessions. The panel evaluated 10 sensory attributes: yellowness, elasticity, fracturability, hardness, dryness, chewiness, adhesiveness and after taste, which are described in Table 2. Two attributes, wheat bread aroma and wheat bread taste intensities were discarded, as no systematic variation was found. Results from one assessor were likewise discarded due to low signal to noise ratio on all attributes. TPA and sensory analysis was repeated another day on a second thawed bread. TPA results were averaged. Replicate variation as well as assessor variation in the use of scale was removed by ANOVA-PLS (A-PLS) with **X** as replicates and assessors 1 to 10 assigned with either zeroes or

ones and **Y** as sensory data. The **Y** residual after 11 PCs was extracted from the A-PLS result, averaged over judges and replicates and was hereafter used as level corrected sensory data in the subsequent analysis (Martens et al., 2000).

Data processing

A total of fifty-two parameters (Table 2) from grain to brain were recorded for twenty-nine wheat lots and arranged in one data matrix (29 × 52). Similarities and dissimilarities of the samples as well as the co-variation of variables were evaluated by principal component analysis (PCA) on mean centred and standardised data in LatentiX (Version 1.00, Latent5, Copenhagen, Denmark, www.latentix.com). The effect of growth year, location/cultivar and catch crop were evaluated by a three-factor ANOVA on all parameters for the twelve catch crop lots and the effect of growth year, location/cultivar and fractionation were likewise evaluated by a three-factor ANOVA on all parameters for the twelve fractionated lots in The Unscrambler (Version 9.2, Camo, Norway). In order to extract and compare the effects of catch crop treatments and TriQ fractionation only, the growth year and location/cultivar variation was removed by orthogonalisation against **Y,** a dummy matrix of 27 samples × 4 groups (03AaVi, 04AaVi, 03KiCo and 04KiCo) with ones and zeros corresponding to 'belong to group' and 'not in group' respectively (Andersson, 1999; Svenstrup et al., 2005).

$$\mathbf{X}_{\text{ortho}} = \mathbf{X} - \mathbf{Y}(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{X} \qquad \text{(Eq. 3)}$$

X was the data matrix with all measurements, except for winter cultivar lots 04AaPent and 04KiCapo which did not belong to any of the groups. Both matrices were mean centred before orthogonalisation and the inverse term in Eq. 3 was in practise the pseudo-inverse as the centred **Y'Y** is not full rank.

Orthogonalisation was performed in Matlab (Version 6.5.0.180913a Release 13, The MathWorks, Inc., USA).

Table 2: List and ranges of quality parameters measured with short descriptions and four letter abbreviations used in Table 3 and Figs. 1B and 2B. Abbreviations: first two letters refers to parameter and last two to origin: ~ = NIT prediction, _ = original test, G = grain, F = flour, FB = fresh bread, TB = thawed bread.

Parameter	Code	Description	Min	Max	Mean	Std.
Predicted by NIT on grain in Foss Infratec 1241						
Protein	Pr~G	Dry matter protein concentration in grain [%]	9.4	13.3	11.4	0.9
Density	De_G	Density of bulk grains determined by weight module [kg/hL]	71.2	86.1	79.2	4.0
Moisture	Mo~G	Moisture content in grain [%]	11.4	16.6	13.4	1.4
Zeleni	Ze~G	Sedimentation test at 15% moisture in grain [mL]	28.7	45.7	35.7	4.4
Starch	Sc~G	Dry matter starch concentration in grain [%]	66.2	69.4	67.7	0.9
W. Gluten	WG~G	Wet gluten in grain determined at 14% moisture [%]	18.2	29.8	23.3	3.5
Predicted by NIT on flour in Foss Infratec 1241						
Moisture	Mo~F	Moisture content in flour [%]	13.9	15.8	14.7	0.5
Ash	As~F	Ash content in flour [%]	0.47	0.61	0.51	0.03
Protein	Pr~F	Dry matter protein concentration in flour [%]	8.8	12.0	10.4	0.8
Water Abs.	WA~F	Farinograph water absorption in flour determined at 14% moisture [%]	51.7	59.1	54.6	2.3
Stability	St~F	Farinograph stability in flour determined at 14% moisture [min]	2.7	8.3	3.6	1.5
W. Gluten	WG~F	Wet gluten in flour determined at 14% moisture [%]	22.8	30.7	26.2	2.1
Standard flour tests						
Falling Num.	FN_F	Falling number determined at 14% moisture [s]	212	442	312	76
Water Abs.	WA_F	Farinograph water absorption in flour determined at 14% moisture [%]	51.1	61.8	56.1	3.3
Dev. Time	DT_F	Farinograph development time in flour determined at 14% moisture [min]	1.3	4.7	2.0	0.8
Stability	St_F	Farinograph stability in flour determined at 14% moisture [min]	0.9	8.0	3.6	2.2
Softening	So_F	Farinograph softening in flour determined at 14% moisture [FU]	45	155	88	25
Zeleni	Ze_F	Sedimentation test at 14% moisture in flour [mL]	26.0	37.5	31.5	2.9
Gel. Temp.	GT_F	Amylograph gelatinisation temperature in the gel. maximum [ºC]	70.4	89.4	80.4	7.7
Gel. Max.	GM_F	Amylograph gelatinisation maximum [AU]	195	1630	685	513
W. Gluten	WG_F	Wet gluten in grain determined at 14% moisture [%]	14.8	27.5	21.3	3.0
Gluten Ind.	GI_F	Gluten quality index, percentage strong gluten [%]	71.2	100.0	95.2	7.5
Bread proportions						
Mass	Ma_B	Mass of bread after cool down [g]	855	919	884	19
W. Loss	WL_B	Relative water loss during baking and cool down [%]	0.19	0.27	0.23	0.02
Volume	Vo_B	Bread volume estimated from circumference after cool down [L]	2.62	3.87	3.19	0.41
Density	De_B	Based om volume and mass after cool down [g/L]	0.23	0.34	0.28	0.04
Texture profil analysis on fresh bread using TA-XT2 Texture Analyser						
Hardness 1	H1FB	Force recorded at 60% compression during first down stroke [N]	3.2	14.5	7.2	3.0
Hardness 2	H2FB	Force recorded at 60% compression during second down stroke [N]	2.4	9.1	4.9	1.9
Adh. Force	AFFB	Maximum negative force recorded during first up stroke [N]	0.02	0.46	0.16	0.10
Cohesiveness	CoFB	Work done during second down stroke relative to work done during first down stroke [-]	0.58	0.73	0.65	0.04
Springiness	SpFB	Recovery height relative to compression length [-]	0.84	0.96	0.90	0.03
Gumminess	GuFB	Gumminess = Cohesiveness × Hardness 1 [N]	2.2	8.4	4.6	1.7
Chewiness	ChFB	Chewiness = Gumminess × Springiness [N]	2.0	7.5	4.1	1.5
Resilience	ReFB	Positive work done during first upstroke rel. to work done during first down stroke [-]	0.22	0.36	0.27	0.03
Adhesiveness	AdFB	Negative work done during first upstroke [Ns]	0.02	2.16	0.54	0.49
Texture profil analysis on thawed bread using Stevens QTS-25						
Hardness 1	H1TB	Maximum force recorded during first down stroke (75% comp) [N]	16.3	61.6	31.9	12.0
Hardness 2	H2TB	Maximum force recorded during second down stroke (75% comp) [N]	13.2	43.5	24.7	8.8
Adh. Force	AFTB	Maximum negative force recorded during first up stroke [N]	0.00	0.12	0.04	0.03
Cohesiveness	CoTB	Work done during second down stroke relative to work done during first down stroke [-]	0.27	0.45	0.35	0.04
Springiness	SpTB	Recovery height relative to compression length [-]	0.74	0.96	0.85	0.05
Gumminess	GuTB	Gumminess = Cohesiveness × Hardness 1 [N]	5.5	21.0	11.0	4.0
Chewiness	ChTB	Chewiness = Gumminess × Springiness [N]	5.0	17.3	9.2	3.3
Resilience	ReTB	Positive work done during first upstroke rel. to work done during first down stroke [-]	0.01	0.05	0.02	0.01
Adhesiveness	AdTB	Negative work done during first upstroke [Ns]	0.01	0.86	0.22	0.23
Sensory panel evaluations on thawed bread						
Yellowness	YeTB	Rating of the colour yellow in the crumb	-4.8	5.1	0.4	3.1
Elasticity	ElTB	Recovery of crumb height rated after compression between thumb and forefinger	-3.5	3.5	0.2	2.0
Fracturability	FrTB	The tendensy of the crumb to fracture while stretching the crumb between the hands	-2.1	4.1	0.4	1.7
Hardness	HaTB	The hardness of crumb rated at first bite	-2.9	2.3	-0.2	1.7
Dryness	DrTB	Rating of the crumb dryness in the mouth when chewing	-3.1	2.8	0.0	1.3
Chewiness	CwTB	The amount of mastications before the crumb bite is ready to swallow	-2.2	1.5	-0.2	1.1
Adhesiveness	AhTB	The tendency of the crumb to stick in the mouth and forming a lump when chewing	-1.9	1.5	-0.1	0.9
Aftertaste	AfTB	Intensity of non-specific after-taste (e.g. sourish, staled, yeast)	-1.0	2.9	0.1	1.3

Results and Discussion

The PCA overview in Figure 1 of the first two components explaining 34% and 21% of the variation by screening the twenty-nine wheat lots with fifty-two parameters shows the most pronounced patterns in the data. In the score plot, Fig. 1A, four distinct groups of lots are seen corresponding to the four combinations of growth year and location/cultivar. First principal component (PC) is thus primarily determined by growth year while second PC is determined by location/cultivar, hence climate, environment and genetics are the fundamental sources of variability with regard to the wheat lots and parameters measured here. As Combi was exclusively grown in Kiel and Vinjett only grown in Aarslev, the effects of location and cultivar are confounded and cannot be distinguished further. However, in the present study the effects of sorting as compared to catch crop diversification is in focus and thus the fundamental variability due to genetics, climate and environment as seen in Fig. 1A conveniently spans the data space. Within each group the various treatments are indicated. Since fractions 1 to 3 (F1, F2 and F3) are sorted from a mixture (Mx) of the field treatments (NC, FR and WV), the fractioned lots within each group was a priory expected to posses properties in the same order of magnitude as the field treatments only spanned systematically according to internal complex quality trait and the variability present in the mixture. The PCA elegantly demonstrates that bulk diversified fractions 1 to 3 span the entire variation range in all groups, fractions 1 and 3 being the extremes. However, in group 03AaVi, the no catch crop (NC) treatment lot appears to be a mild outlier. Thorough inspection of data indicates that the data for this particular lot are indeed generally relatively extreme compared to the data of lots from the same year and location/cultivar. Although extreme relative to the group, it does not represent an unlikely variation. At this stage, it is not detrimental for the analysis. The direction of the fractions 1 to 3 appears to be influenced by both PC1 and PC2,

thus the effect of fractionation appears to be systematic as opposed to the catch crop treatment. The winter cultivars 04KiCapo and 04AaPent are apparently quite similar to the Combi lots from Kiel 2004 (04Ki).

The corresponding loadings for the fifty-two parameters in Fig. 1B are coherently clustered and should be examined along with the two ANOVAs in Table 3. The overall confirmatory coherence between the parameters measured at all levels of the process from grain to brain validates the results of individual parameters. Most of the significant parameters for factor year have high absolute scores on PC1 and likewise for the significant parameters for location/cultivar on PC2. However the PCA plot (Fig. 1A and B) contains more than year and location/cultivar variation in the first two PC's and may appear messy without the statistical information in Table 3. The block effects of growth year and location/cultivar are not of particulate interest here, however included for completeness and only briefly covered in the following. Bread texture and volume along with enzyme activity and gluten content and quality span the first PC (Fig. 1B). This immediate major quality differentiation by physical appearance where lots from 2003 are more compact than lots from 2004 can seen by inspecting Fig 1A and 1B simultaneously. This is probably primarily due to higher *alpha*-amylase activity in the more recent lots (2004) represented by falling number (FN_F), gelatinisation temperature (GT_F) and gelatinisation maximum (GM_F) in cluster 3 in Fig 1B. The falling number for 2003 is recorded to 329 – 442 s, while for 2004 it is 212 – 318 s. Correspondingly the gelatinisation maximum and gelatinisation temperature varies from 545 to 1630 and from 86.1 ºC to 89.4 ºC respectively in 2003 lots, while for 2004 they are varying from 195 to 450 and from 70.4 ºC to 82.2 ºC respectively. This is confirmed in Table 3 in which all *alpha*-amylase activity measures as well as bread volume, density and hardness are highly significant ($P < 0.001$ and $P <$

0.01) with respect to the growth year factor. The potential contrary effect on bread volume due to slightly higher wet gluten content in 2003 compared to 2004 ($P < 0.01$) was probably masked by the strong difference in enzymatic activity level. The location/cultivar (PC2, Fig. 1B) affected primarily Farinograph water absorption ($P < 0.001$) with pronounced effects on bread mass and related texture parameters, such as cohesiveness of fresh bread ($P < 0.01$) and sensory elasticity and dryness attributes ($P < 0.01$ and $P < 0.001$).

Going through the effects of varying catch crops in Table 3, it is clear that this strategy as a mean to diversify crop quality needs more attention in the choice of catch crop or the procedures for optimising the effect. No systematic, coherent and significant diversification was observed. However catch crops may still be a sensible way to increase yield and reduce nutrient loss and the resulting environmental impact (Eriksen et al., 2006; Francis et al., 1992; Kristensen and Thorup-Kristensen, 2004). The apparent significant results and results with low P value with respect to cohesiveness, springiness and resilience in TPA of fresh bread and hardness 1 and 2, gumminess and chewiness in TPA of thawed bread should not be considered viable results. They lack coherence with other physicochemical parameters and are redundant in each block of data. CoFB, SpFB and ReFB are describing more or less the same spongy or elasticity property in the fresh bread, which is not confirmed in thawed bread, where H1TB and H2TB are naturally correlated to GuTB and ChTB as they are functions of H1TB (Table 2). So even though five $P < 0.05$ (nearly six) indicate five (six) features of interest they are only due to two underlying phenomena which properly appear significant due random noise in the already rather noisy data from TPA.

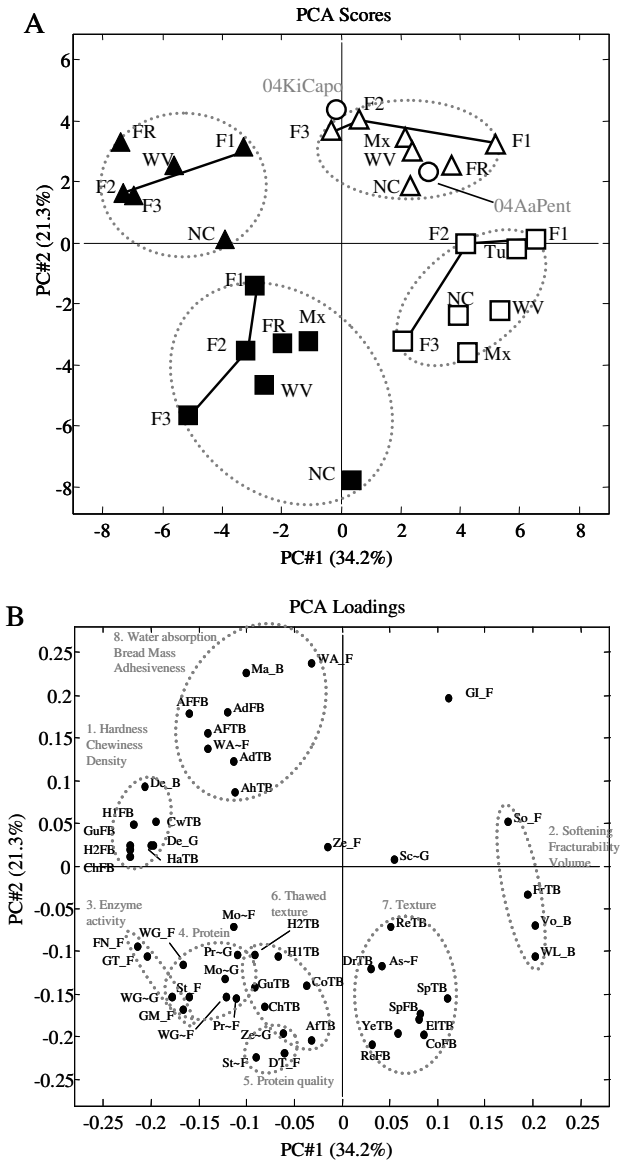


Figure 1: Principal component analysis (PCA) of twenty-nine samples and fifty-two variables. A: Score plot of 03AaVi (■), 03KiCo (▲),04AaVi (□), 04KiCo (△),04KiCapo and 04AaPent (○). PC1 explains primarily the year variation and PC2 primarily the location/cultivar variation. The two winter cultivars, 04KiCapo and 04AaPent are located in the same group as 04Ki. TriQ fractions 1, 2 and 3 are connected within each group. B: Loadings plot with all parameters arranged in coherent clusters especially important for year and location/cultivar variation.

Table 3: Two separate three factor ANOVAs for every quality parameter. 1. Catch crop treatment. 2. TriQ SKNIR sorting fractionation. Year and location/cultivar are blocks in each ANOVA.

		Catch crop ANOVA			Fraction ANOVA		
Data block	Attribute	Year	Location	Treatment	Year	Location	Treatment
NIT on grain	Pr~G	0.12 ^{NS}	0.83 ^{NS}	0.13 ^{NS}	0.085 ^{NS}	0.26 ^{NS}	0.0023 **
	De_G	0.000010 ***	0.0015 **	0.64 ^{NS}	0.00000000020 ***	0.000000043 ***	0.000080 ***
	Mo~G	0.0017 **	0.071 ^{NS}	0.76 ^{NS}	0.15 ^{NS}	0.84 ^{NS}	1.00 ^{NS}
	Ze~G	0.36 ^{NS}	0.26 ^{NS}	0.16 ^{NS}	0.11 ^{NS}	0.25 ^{NS}	0.0028 **
	Sc~G	0.39 ^{NS}	0.88 ^{NS}	0.57 ^{NS}	0.89 ^{NS}	0.012 *	0.032 *
	WG~G	0.000032 ***	0.12 ^{NS}	0.16 ^{NS}	0.00023 ***	0.99 ^{NS}	0.0018 **
NIT on flour	Mo~F	0.13 ^{NS}	0.63 ^{NS}	0.85 ^{NS}	0.10 ^{NS}	0.60 ^{NS}	0.94 ^{NS}
	As~F	0.96 ^{NS}	0.0082 **	0.72 ^{NS}	0.41 ^{NS}	0.45 ^{NS}	0.13 ^{NS}
	Pr~F	0.051 ^{NS}	0.50 ^{NS}	0.14 ^{NS}	0.073 ^{NS}	1.00 ^{NS}	0.0024 **
	WA~F	0.017 *	0.0043 **	0.67 ^{NS}	0.15 ^{NS}	0.0054 **	0.22 ^{NS}
	St~F	0.013 *	0.054 ^{NS}	0.75 ^{NS}	0.027 *	0.085 ^{NS}	0.47 ^{NS}
	WG~F	0.048 *	0.56 ^{NS}	0.17 ^{NS}	0.032 *	0.37 ^{NS}	0.0025 **
Flour tests	FN_F	0.000026 ***	0.91 ^{NS}	0.86 ^{NS}	0.0000019 ***	0.72 ^{NS}	0.010 *
	WA_F	0.23 ^{NS}	0.0000087 ***	0.30 ^{NS}	0.018 *	0.00000090 ***	0.010 *
	DT_F	0.12 ^{NS}	0.23 ^{NS}	0.41 ^{NS}	0.11 ^{NS}	0.052 ^{NS}	0.078 ^{NS}
	St_F	0.00021 ***	0.35 ^{NS}	0.59 ^{NS}	0.051 ^{NS}	0.46 ^{NS}	0.20 ^{NS}
	So_F	0.064 ^{NS}	0.24 ^{NS}	0.90 ^{NS}	0.021 *	0.64 ^{NS}	0.050 *
	Ze_F	0.58 ^{NS}	0.41 ^{NS}	0.62 ^{NS}	0.48 ^{NS}	0.61 ^{NS}	0.051 ^{NS}
	GT_F	0.00000028 ***	0.98 ^{NS}	0.58 ^{NS}	0.0000000065 ***	0.25 ^{NS}	0.046 *
	GM_F	0.00026 ***	0.078 ^{NS}	0.59 ^{NS}	0.00037 ***	0.041 *	0.69 ^{NS}
	WG_F	0.0092 **	0.78 ^{NS}	0.23 ^{NS}	0.0024 **	0.44 ^{NS}	0.00094 ***
	GI_F	0.022 *	0.22 ^{NS}	0.45 ^{NS}	0.010 *	0.065 ^{NS}	0.50 ^{NS}
Bread dim.	Ma_B	0.035 *	0.0000047 ***	0.90 ^{NS}	0.050 ^{NS}	0.00000066 ***	0.0037 **
	WL_B	0.00062 ***	0.00062 ***	0.31 ^{NS}	0.0024 **	0.045 *	0.13 ^{NS}
	Vo_B	0.0078 **	0.045 *	0.56 ^{NS}	0.0042 **	0.21 ^{NS}	0.36 ^{NS}
	De_B	0.0042 **	0.013 *	0.58 ^{NS}	0.0021 **	0.054 ^{NS}	0.35 ^{NS}
Fresh bread	H1FB	0.0065 **	0.12 ^{NS}	0.60 ^{NS}	0.00011 ***	0.0076 **	0.13 ^{NS}
TPA	H2FB	0.0028 **	0.18 ^{NS}	0.66 ^{NS}	0.000068 ***	0.023 *	0.13 ^{NS}
	AFFB	0.19 ^{NS}	0.020 *	0.49 ^{NS}	0.0077 **	0.0018 **	0.70 ^{NS}
	CoFB	0.95 ^{NS}	0.0047 **	0.046 *	0.79 ^{NS}	0.0069 **	0.38 ^{NS}
	SpFB	0.35 ^{NS}	0.024 *	0.085 ^{NS}	0.17 ^{NS}	0.052 ^{NS}	0.057 ^{NS}
	GuFB	0.0040 **	0.18 ^{NS}	0.74 ^{NS}	0.000028 ***	0.011 *	0.081 ^{NS}
	ChFB	0.0040 **	0.25 ^{NS}	0.84 ^{NS}	0.000019 ***	0.011 *	0.046 *
	ReFB	0.31 ^{NS}	0.0051 **	0.041 *	0.25 ^{NS}	0.010 **	0.075 ^{NS}
	AdFB	0.30 ^{NS}	0.025 *	0.38 ^{NS}	0.27 ^{NS}	0.0022 **	0.68 ^{NS}
Thawed bread	H1TB	0.24 ^{NS}	0.83 ^{NS}	0.049 *	0.072 ^{NS}	0.31 ^{NS}	0.87 ^{NS}
TPA	H2TB	0.15 ^{NS}	0.99 ^{NS}	0.047 *	0.037 *	0.36 ^{NS}	0.93 ^{NS}
	AFTB	0.27 ^{NS}	0.12 ^{NS}	0.32 ^{NS}	0.019 *	0.0095 **	0.40 ^{NS}
	CoTB	0.22 ^{NS}	0.52 ^{NS}	0.26 ^{NS}	0.38 ^{NS}	0.40 ^{NS}	0.043 *
	SpTB	0.39 ^{NS}	0.0048 **	0.32 ^{NS}	0.31 ^{NS}	0.31 ^{NS}	0.97 ^{NS}
	GuTB	0.063 ^{NS}	0.77 ^{NS}	0.026 *	0.023 *	0.20 ^{NS}	0.85 ^{NS}
	ChTB	0.062 ^{NS}	0.35 ^{NS}	0.016 *	0.025 *	0.14 ^{NS}	0.76 ^{NS}
	ReTB	0.40 ^{NS}	0.88 ^{NS}	0.60 ^{NS}	0.79 ^{NS}	0.34 ^{NS}	0.90 ^{NS}
	AdTB	0.68 ^{NS}	0.45 ^{NS}	0.40 ^{NS}	0.10 ^{NS}	0.027 *	0.45 ^{NS}
Thawed bread	YeTB	0.80 ^{NS}	0.00080 ***	0.80 ^{NS}	0.77 ^{NS}	0.00016 ***	0.81 ^{NS}
Sensory eval.	ElTB	0.12 ^{NS}	0.00017 ***	0.68 ^{NS}	0.95 ^{NS}	0.00021 ***	0.11 ^{NS}
	FrTB	0.041 *	0.047 *	0.94 ^{NS}	0.00080 ***	0.79 ^{NS}	0.052 ^{NS}
	HaTB	0.058 ^{NS}	0.33 ^{NS}	0.95 ^{NS}	0.0017 **	0.23 ^{NS}	0.18 ^{NS}
	DrTB	0.69 ^{NS}	0.0068 **	0.28 ^{NS}	0.12 ^{NS}	0.00076 ***	0.00041 ***
	CwTB	0.065 ^{NS}	0.10 ^{NS}	0.80 ^{NS}	0.0022 **	0.32 ^{NS}	0.50 ^{NS}
	AhTB	0.25 ^{NS}	0.034 *	0.87 ^{NS}	0.14 ^{NS}	0.35 ^{NS}	0.14 ^{NS}
	AfTB	0.095 ^{NS}	0.032 *	0.96 ^{NS}	0.13 ^{NS}	0.046 *	0.56 ^{NS}

Fractionation on the other hand clearly exhibit pronounced significant effects for sixteen quality parameters. In addition fifteen quality parameters have close to significant P-values and at the same time lower P-values than for the corresponding catch crop factor (Table. 3). This was further visualised by removing the year and location/cultivar variation by employing orthogonalisation (Eq. 3) to the data matrix excluding lots 04AaPent and 04KiCapo which did not belong to any of the groups. Now only the effects of catch crop treatments and fractionation remain. By PCA of the othogonalized matrix without 03AaViNC, which in this context appeared to be an outlier, a clear treatment pattern emerged, which may be viewed in Fig 2 parallel to Table 3, last column. The scores plot in Fig. 2A shows that the first PC separates the fractions and at the same time show that the sorting diversification expanded the fundamental variation in the original lots encapsulated by the circle. It is noteworthy that the low quality fractions (F1s) are differentiated along PC2, while higher quality fractions, F2 and F3 are more specific. This is conceptually in agreement with the two component unsupervised calibration model which controls the sorting (Löfqvist and Nielsen, 2003). The significant parameters (Table 3) as well as those non-significant but low P values are responsible for the major variation along the first PC, indicated by '*' and 'o' in Fig. 2B. The short distance in the loadings plot (Fig. 2B) between the related parameters measured at various steps in the process validates the findings. Protein contents determined in grain and flour (Pr~G and Pr~F) are present in the exact same spot, thus confirms the findings. Likewise for wet gluten contents determined from grain NIT spectra (WG~G), flour NIT spectra (WG~F) and standard test (WG_F). Zeleny sedimentation volume (Ze~G and Ze_F) and Farinograph water absorption (WA~F and WA_F) also have similar values along the first PC regardless of how they were determined. None of the residual variation at higher order PCs (not shown) was found to explain variation due to fractionation. The

preceding catch crops used in the field did not have any systematic effect on the grouping of samples in the PCA scores plot (Fig. 2A) and confirm the findings in the ANOVA (Table 3). Thus, it is observed that the protein content, the quality of protein in terms of wet gluten and Zeleny sedimentation value and the subsequent dough handling properties in the Farinograph and the *alpha*-amylase activity are all differentiated by the TriQ fractionation. The fractionation strategy thus works well for these quality parameters on the different initial mixtures, with fraction 1 as low, fraction 2 as medium and fraction 3 as high quality. Having a high falling number may not be beneficial in itself as it leads to compact bread, but by keeping it relatively high; the possibility to regulate the amylase activity level is preserved.

The resulting breads are consequently differentiated on mass ($P < 0.01$) and sensory perceived dryness ($P < 0.001$) due to water absorption ($P < 0.05$) (Table 3). TPA chewiness of the fresh bread and cohesiveness of the thawed bread are both significant, however only with $P < 0.05$. The apparent lack of significant, coherent parameters regarding bread quality due to fractionation is most probably due to an unfortunate baking test method (see below).

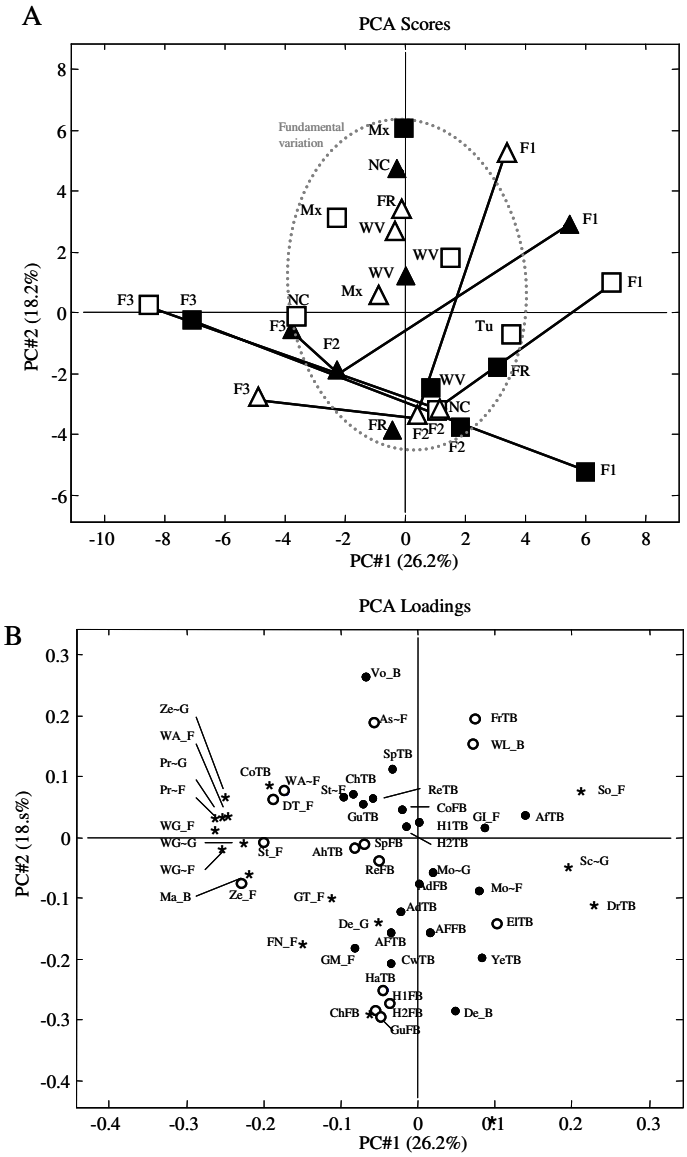


Figure 2: Principal component analysis (PCA) of twenty-six samples (without winter cultivars 04AaPent and 04KiCapo and outlier 03AaViNC) after removal of year and location/cultivar variations by orthogonalization (Eq. 3). A: Scores plot of 03AaVi (■), 03KiCo (▲),04AaVi (□) and 04KiCo (△) where PC1 explains the major variance in TriQ fractions which are connected with lines in each group. B: The loadings plot with ANOVA significant parameters (*) and parameters with low but non-significant P value for fractionation (O).

In summery the raw data of central grain and flour quality parameters extracted by the preceding multivariate analysis along with bread volume are presented in Table 4 for Aarslev/Vinjett lots and in Table 5 for Kiel/Combi lots. The data include results from catch crop treatments and their weighted means

corresponding to the mixture value and results from SKNIR fractionation and their weighted means. The means are included for validation of the results of individual fractions. The actual diversification obtained by fractionation can be compared to the variation in the starting material in terms of the different catch crop treatments. It is apparent that the systematic effect of SKNIR fractionation is pronounced, not only by significance level, but also by actual levels of the individual parameters. From the figures in columns F1, F2 and F3 in Table 4 and 5 it is evident, that the concept of sorting may add value to an otherwise low quality crop. In the lots used here the flour protein was increased by 0.5 to 1.7 %-point, wet gluten by 1.8 to 5.5 %-point and Zeleny sedimentation value by 1.4 to 3.5 mL as compared to the weighted means of the starting material. Comparable but higher values were found at the grain level. Flour water absorption was increased by 0.5 to 1.4 %-point, development time by 0 to 1.6 min, stability by -0.3 to 4.0 min, softening decrease by 6 to 34 FU and falling number increase by 10 to 48 s in fraction 3 as compared to the weighted mean of the starting material. Thus in this experiment, one third of the starting material was increased in quality and value (Fraction 3), one third remained close to average quality (Fraction 2) and one third was left of low baking quality (Fraction 1). The included farinograms gives a clear graphical representation of the dough consistency characteristics in the three fractions with increasing development time and stability through the fraction from 1 to 3.

The obtainable differentiation and level of quality by TriQ sorting is only determined by the inherent heterogeneity of the lot being fractioned and the target size of the fractions. The differentiation between low (F1) and high (F3) quality regarding protein, wet gluten and water absorption for instance is greater for the Aarslev lots as compared to the Kiel lots. The heterogeneity of these parameters was thus bigger in Aarslev prior to the sorting. Larger differentiation

is obtainable if the fraction size is not an issue, but rather a specific target quality is the aim of the fractionation regardless of the potential yield (Nielsen, 2002). The potential increased value of fractioning a given lot should be based on a test on a representative sample to evaluate the heterogeneity (Tønning et al., 2006) and the potential outcome of the sorting.

The expected increasing bread volume with increasing fraction number, hence increasing protein content and quality (Bushuk et al., 1969), is only seen in the Aarslev/Vinjett lots, where the difference in protein content and wet gluten content between fraction 1 and fraction 3 is much larger than in the Kiel/Combi lots. Counteracting interaction between enzyme activity levels and protein content and quality, and especially an unfortunate choice of baking test not capable of differentiating wheat lots (Grausguber et al., 2001; Zwingelberg and Brümmer, 1990) appears to mask the potential effect of the SKNIR diversified protein quality on bread volume and further bread quality parameters. Although the SD-253 provided reproducible loaf volumes with relative standard deviation < 3% similar to other reports (Dubuc and Boudreau, 1992; Hansen and Hansen, 1992, 1993; Peltonen and Salovaara, 1991), the kneading procedure in the automatic home-bakery apparently favours weak flours to ensure relatively good baking results regardless of the flour quality used. The baking test would probably provide much better results if the mixing was conducted outside the baking machine (Grausgruber et al., 2001; Zwingelberg and Brümmer, 1990), e.g. in the Farinograph mixing chamber (Peltonen and Salovaara, 1991).

Table 4: Raw data of parameters of particular interest and farinograms for field treatments as well as SKNIR fractionation for Aarslev Vinjett lots. ***, ** and * refers to significance levels P < 0.001, 0.01 and 0.05 respectively in separate three factor ANOVAs for every quality parameter with respect to SKNIR fractionation with year and location/cultivar as blocks from Table 3. *Italic* numbers are weighted means of three field treatments and three SKNIR fractions. **Bold** numbers are systematically affected by SKNIR fractionation.

		Abr.	Units	Catch crop treatment				SKNIR fractionation			
				NC	FR/Tu	WV	*W. Mean*	F1	F2	F3	*W. Mean*
■ **2003 Aarslev Vinjett (03AaVi)**											
Grain	Density ***	De_G	[kg/hL]	82.1	80.2	82.1	*81.4*	**80.2**	**81.1**	**81.3**	*80.9*
	Protein **	Pr~G	[%]	13.3	11.0	11.4	*11.6*	**10.3**	**11.3**	**12.7**	*11.5*
	W. Gluten ***	WG~G	[%]	29.8	25.0	26.3	*26.6*	**23.8**	**26.1**	**28.8**	*26.4*
	Zeleny **	Ze~G	[mL]	45.7	33.3	35.9	*37.0*	**30.0**	**36.3**	**44.3**	*37.3*
Flour	Protein **	Pr~F	[%]	12.0	10.1	10.4	*10.6*	**9.5**	**10.4**	**11.9**	*10.7*
	W. Gluten ***	WG_F	[%]	27.5	20.8	21.8	*22.6*	**20.2**	**21.3**	**26.5**	*22.8*
	Gluten Ind.	GI_F	[%]	71.2	89.2	90.6	*86.1*	85.7	86.1	82.1	84.5
	Zeleny	Ze_F	[mL]	30.9	27.4	29.5	*29.1*	**26.0**	**28.8**	**30.5**	*28.6*
	Fall. Num *	FN_F	[s]	376	399	397	*393*	**372**	**404**	**404**	*394*
	Gel. Temp. *	GT_F	[°C]	87.6	88.1	89.1	*88.4*	**89.1**	**88.7**	**88.7**	*88.8*
	Gel. Max.	GM_F	[AU]	1075	1630	1595	*1499*	1500	1540	1530	*1524*
	Water Abs. *	WA_F	[%]	54.4	51.7	52.1	*52.4*	**51.1**	**51.5**	**53.5**	*52.1*
	Dev. Time	DT_F	[min]	4.7	1.9	2.0	*2.5*	**1.9**	**2.3**	**4.6**	*3.0*
	Stability	St_F	[min]	5.2	4.8	5.4	*5.2*	**4.4**	**5.6**	**4.9**	*5.0*
	Softening *	So_F	[FU]	105	95	80	*90*	**80**	**80**	**70**	*76*
Bread	Volume	Vo_B	[L]	3.67	2.87	2.88	*3.04*	2.61	2.74	3.08	2.82
□ **2004 Aarslev Vinjett (04AaVi)**											
Grain	Density ***	De_G	[kg/hL]	76.6	76.0	75.8	*76.1*	**74.3**	**75.2**	**76.0**	*75.2*
	Protein **	Pr~G	[%]	12.4	10.4	11.1	*11.2*	**9.4**	**10.7**	**12.7**	*10.9*
	W. Gluten ***	WG~G	[%]	22.2	19.6	21.0	*20.8*	**19.5**	**20.6**	**25.9**	*21.9*
	Zeleny **	Ze~G	[mL]	41.7	33.6	37.6	*37.3*	**28.9**	**34.4**	**43.6**	*35.4*
Flour	Protein **	Pr~F	[%]	11.0	9.6	10.1	*10.2*	**8.8**	**10.0**	**11.9**	*10.2*
	W. Gluten ***	WG_F	[%]	22.7	17.5	19.2	*19.5*	**14.8**	**18.6**	**25.0**	*19.3*
	Gluten Ind.	GI_F	[%]	100.0	100.0	100.0	*100.0*	100.0	100.0	100.0	*100.0*
	Zeleny	Ze_F	[mL]	34.6	31.1	32.8	*32.7*	**28.5**	**36.5**	**37.5**	*34.0*
	Fall. Num *	FN_F	[s]	252	234	234	*239*	**212**	**237**	**250**	*232*
	Gel. Temp. *	GT_F	[°C]	74.6	73.2	72.3	*73.3*	**71.3**	**72.0**	**73.7**	*72.3*
	Gel. Max.	GM_F	[AU]	335	280	260	*290*	195	245	290	*242*
	Water Abs. *	WA_F	[%]	54.7	52.9	53.4	*53.6*	**52.5**	**53.6**	**55.1**	*53.7*
	Dev. Time	DT_F	[min]	2.1	1.8	1.9	*1.9*	**1.6**	**1.7**	**2.5**	*1.9*
	Stability	St_F	[min]	3.9	1.8	2.2	*2.5*	**1.7**	**2.4**	**8.0**	*4.0*
	Softening *	So_F	[FU]	80	105	105	*98*	**125**	**85**	**55**	*89*
Bread	Volume	Vo_B	[L]	3.85	3.71	3.58	*3.71*	3.53	3.29	3.87	*3.56*

Table 5: Raw data of parameters of particular interest and farinograms for field treatments as well as SKNIR fractionation for Kiel Combi lots. ***, ** and * refers to significance levels P < 0.001, 0.01 and 0.05 respectively in separate three factor ANOVAs for every quality parameter with respect to SKNIR fractionation with year and location/cultivar as blocks from Table 3. *Italic* numbers are weighted (W) means of three field treatments and three SKNIR fractions. **Bold** numbers are systematically affected by SKNIR fractionation.

		Abr.	Units	Catch crop treatment NC	FR/Tu	WV	*W. Mean*	SKNIR fractionation F1	F2	F3	*W. Mean*
▲ **2003 Kiel Combi (03KiCo)**											
Grain	Density ***	De_G	[kg/hL]	86.1	85.8	85.7	*85.9*	**82.9**	**83.4**	**83.9**	*83.4*
	Protein **	Pr-G	[%]	12.4	11.9	11.9	*12.1*	**11.3**	**12.2**	**12.2**	*11.9*
	W. Gluten ***	WG-G	[%]	27.4	26.3	26.2	*26.6*	**24.6**	**26.9**	**27.1**	*26.3*
	Zeleny **	Ze-G	[mL]	39.2	35.9	36.5	*37.2*	**33.9**	**37.1**	**38.4**	*36.5*
Flour	Protein **	Pr-F	[%]	11.0	10.5	10.7	*10.7*	**10.0**	**11.2**	**11.2**	*10.8*
	W. Gluten ***	WG_F	[%]	24.6	23.9	24.5	*24.3*	**20.9**	**24.6**	**26.1**	*23.9*
	Gluten Ind.	GI_F	[%]	94.1	87.7	100.0	*94.0*	100.0	100.0	89.6	*96.6*
	Zeleny	Ze_F	[mL]	34.1	33.1	32.5	*33.2*	**30.2**	**35.2**	**36.0**	*33.9*
	Fall. Num *	FN_F	[s]	343	423	395	*387*	**329**	**408**	**442**	*394*
	Gel. Temp. *	GT_F	[°C]	87.6	89.3	88.4	*88.4*	**86.1**	**89.1**	**89.4**	*88.3*
	Gel. Max.	GM_F	[AU]	785	1065	890	*914*	545	940	935	*813*
	Water Abs. *	WA_F	[%]	58.4	58.5	58.4	*58.4*	**57.5**	**58.7**	**58.9**	*58.4*
	Dev. Time	DT_F	[min]	2.1	2.0	2.3	*2.1*	**1.5**	**1.7**	**2.1**	*1.8*
	Stability	St_F	[min]	5.9	5.5	4.8	*5.4*	**3.2**	**7.1**	**7.3**	*5.9*
	Softening *	So_F	[FU]	60	50	65	*58*	**80**	**50**	**45**	*58*
Bread	Volume	Vo_B	[L]	2.83	2.66	2.83	*2.78*	2.81	2.64	2.67	*2.70*
▲ **2004 Kiel Combi (04KiCo)**											
Grain	Density ***	De_G	[kg/hL]	77.7	77.7	77.9	*77.8*	**77.2**	**78.0**	**78.4**	*77.9*
	Protein **	Pr-G	[%]	10.9	11.5	11.4	*11.3*	**10.5**	**11.0**	**12.1**	*11.2*
	W. Gluten ***	WG-G	[%]	18.2	18.9	18.8	*18.7*	**20.5**	**21.3**	**24.2**	*22.0*
	Zeleny **	Ze-G	[mL]	32.2	35.5	34.4	*34.1*	**28.7**	**31.6**	**36.2**	*32.2*
Flour	Protein **	Pr-F	[%]	9.6	10.1	10.0	*9.9*	**9.5**	**9.8**	**10.8**	*10.0*
	W. Gluten ***	WG_F	[%]	18.3	20.1	20.2	*19.6*	**17.6**	**19.0**	**22.1**	*19.6*
	Gluten Ind.	GI_F	[%]	100.0	100.0	100.0	*100.0*	100.0	100.0	100.0	*100.0*
	Zeleny	Ze_F	[mL]	30.4	32.2	30.8	*31.2*	**27.3**	**30.3**	**34.7**	*30.8*
	Fall. Num *	FN_F	[s]	272	228	242	*246*	**221**	**241**	**263**	*242*
	Gel. Temp. *	GT_F	[°C]	75.8	72.0	71.7	*73.1*	**70.4**	**71.7**	**72.9**	*71.7*
	Gel. Max.	GM_F	[AU]	330	230	240	*264*	210	240	260	*237*
	Water Abs. *	WA_F	[%]	58.3	59.0	58.9	*58.8*	**58.6**	**58.3**	**59.7**	*58.8*
	Dev. Time	DT_F	[min]	1.3	1.7	1.5	*1.5*	**1.5**	**1.3**	**1.6**	*1.5*
	Stability	St_F	[min]	1.0	1.8	1.5	*1.5*	**0.9**	**1.5**	**1.6**	*1.3*
	Softening *	So_F	[FU]	120	90	100	*103*	**155**	**110**	**90**	*117*
Bread	Volume	Vo_B	[L]	3.09	3.53	3.37	*3.34*	3.71	2.98	3.03	*3.22*

Conclusions

This paper demonstrates that the fundamental variations in the bulk material is determined by climate (year), location (soil, local climate) and cultivar (genetics) while the SKNIR fractionation on top of this variation base is superior as compared to catch crop diversification. The catch crops may however still be useful in the prevention of recipient pollution and nutrient conservation in the field. While this study demonstrates that the SKNIR sorting may serve as a way to increase crop value in organic farming as compared to various field treatments with insignificant effect its potential reach far beyond. In the future we foresee the sorting associated with appropriate chemometric calibrations as a breeding tool for improved nutritional value and for bioactive compounds such as betain, dietary fibres, polyphenols etc. and in quality added value in specialised high value gourmet products, bread production, biscuit production or animal feed and simultaneously as a filter for defect mycotoxin infected seeds (outliers). The present very high prices on the world market for energy combined with an increasing demand of cereals for feed in South East Asia have recently drastically increased the world market price on wheat. This tendency if sustainable will support industrial SKNIR sorting in the future. Even low quality fractions, e.g. mycotoxin infested fractions, may be economically utilized for heating and bio-ethanol. Emerging prototype sorting equipment with a capacity of 1 to 10 T h^{-1} may very well be integrated in a farm silo system – sorting the crop as it is harvested and stored. Alternatively a mobile system may become a part of the standard equipment in the regional machine pool. By up-scaling the system it may also be used by the milling industry.

Acknowledgements

We are grateful to the careful laboratory work by Jette Ryaa Nielsen and the in-house sensory panel at University of Aarhus, Faculty of Agricultural Sciences, Department of Food Science and Lisbeth T. Hansen at University of Copenhagen, Faculty of Life Sciences, Department of Food Science. Bo Löfqvist from BoMill AB is gratefully acknowledged for conducting fractionation of our starting material and for valuable discussions. We thank Cerealia Mills, Vejle, Denmark for facilitating laboratory help, space and equipment. We are grateful to Johannes Ravn Jørgensen and Bernd Wollenweber at University of Aarhus, Faculty of Agricultural Sciences, Department of Genetics and Biotechnology for help and enlightening discussions. For financial support we are greatly thankful to the Interreg IIIA programme between Fyns Amt, Odense, Denmark and Technologie-Region K.E.R.N. Rendsburg, Germany and to University of Copenhagen, Faculty of Life Sciences for the ph.d. scholarship to the first author, and to the project FFS05-9: Build Your Food sponsored by the Ministry of Food, Agriculture and fisheries.

References

Andersson, C.A., 1999. Direct orthogonalization. Chemometrics and Intelligent Laboratory Systems 47, 51-63.

Armstrong, P.R., 2006. Rapid single-kernel NIR measurement of grain and oil-seed attributes. Applied Engineering in Agriculture 22(5), 767-772.

Bushuk, W., Briggs, K.G., Shebeski, L.H., 1969. Protein quantity and quality as factors in the evaluation of bread wheats. Canadian journal of plant science 49(2), 113-122.

Delwiche, S.R., 1993. Measurement of Single-Kernel Wheat Hardness Using Near-Infrared Transmittance. Transactions of the ASEA 36(5), 1431-1437.

Delwiche, S.R., 1995. Single Wheat Kernel Analysis by Near-Infrared Transmittance - Protein-Content. Cereal Chemistry, 72(1), 11-16.

Delwiche, S.R., 1998. Protein content of single kernels of wheat by near-infrared reflectance spectroscopy. Journal of Cereal Science 27(3), 241-254.

Delwiche, S.R., Graybosch, R.A., Peterson, C.J., 1998. Predicting protein composition, biochemical properties, and dough-handling properties of hard red winter wheat flour by near-infrared reflectance. Cereal Chemistry 75(4), 412-416.

Delwiche, S.R., Hruschka, W.R., 2000. Protein content of bulk wheat from near-infrared reflectance of individual kernels. Cereal Chemistry, 77, 86-88.

Delwiche, S.R., Pearson, T.C., Brabec, D.L., 2005. High-speed optical sorting of soft wheat for reduction of deoxynivalenol. Plant Disease 89(11), 1214-1219.

Delwiche, S.R., Weaver, G., 1994. Bread quality of wheat flour by near-infrared spectrometry: Feasibility of modelling. Journal of Food Science 59(2), 410-415.

Dowell, F.E., Maghirang, E.B., Graybosch, R.A., Baenziger, P.S., Baltensperger, D.D., Hansen, L.E., 2006. An automated near-infrared system for selecting individual kernels based on specific quality characteristics. Cereal Chemistry 83(5), 537-543.

Dubuc, J.P., Boudreau, A., 1992. Prediction of breadmaking quality for wheat breeding by a robotic baking method. Cereal Research Communications 20, 105-110.

Elfverson, C., Andersson, A.A.M., Aman, P., Regener, S., 1999. Chemical composition of barley cultivars fractionated by weighing, pneumatic classification, sieving, and sorting on a specific gravity table. Cereal Chemistry 76(3), 434-438.

Eriksen, J., Pedersen, L., Jørgensen, J.R., 2006. Nitrate leaching and bread making quality of spring wheat following cultivation of different grasslands. Agriculture Ecosystems and Environment 116(3-4), 165-175.

Francis, G.S., Haynes, R.J., Sparling, G.P., Ross, D.J., Williams, P.H., 1992. Nitrogen mineralization, nitrate leaching and crop growth following cultivation of a temporary leguminous pasture in autumn and winter. Fertilizer Research 33(1), 59-70.

Fredriksson, H., Salomonsson, L., Salomonsson, A.-C., 1997. Wheat cultivated with organic fertilisers and urea: Baking performance and dough properties. Acta Agriculturae Scandinavica Section B-Soil and Plant Science 47, 35-42.

Grausgruber, H., Kreuzmayr, A.E., Ruckenbauer, P., 2001. Evaluation of the breadmaking quality of Austrian-grown wheats using an automatic home-bakery. Cereal Research Communications 29(3-4), 421-428.

Haglund, Å., Johansson, L., Dahlstedt, L., 1998. Sensory evaluation of wholemeal bread from ecologically and conventionally grown wheat. Journal of Cereal Science 27(2), 199-207.

Hansen, B., Hansen, Å., 1992. Test baking of bread by household baking machine. Food Science and Technology, 25, 585-587.

Hansen, B., Hansen, Å., 1993. Erratum. Test baking of bread by household baking machine. Food Science and Technology 26(2), 181.

Johansson, E., Prieto-Linde, M.L., Jönsson, J.Ö., 2001. Effects of wheat cultivar and nitrogen application on storage protein composition and breadmaking quality. Cereal Chemistry 78(1), 19-25.

Kihlberg, I., Johansson, L., kohler, A., Risvik, E., 2004. Sensory qualities of whole wheat pan bread – influence of farming system, milling and baking technique. Journal of Cereal Science 39, 67-84.

Kristensen, H.L., Thorup-Kristensen, K., 2004. Root growth and nitrate uptake of three different catch crops in deep soil layers. Soil Science Society of America Journal 68(2), 529-537.

Löfqvist, B., Nielsen, J.P., 2003. A method for sorting objects comprising organic material. World Patent WO 03/004179 A1.

Löfqvist, B., Nielsen, J.P., 2006. Method and device for sorting objects. US Patent 2006/0144762 A1

Martens, M., Bredie, W.L.P., Martens, H., 2000. Sensory profiling data studied by partial least squares regression. Food Quality and Preference 11, 147-149.

Nielsen, J.P., 2002. Fast quality assessment of barley and wheat: Chemometric exploration of instrumental data with single seed applications. Ph.D. thesis. Food Technology, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Frederiksberg, Denmark.

Nielsen, J.P., Pedersen, D.K., Munck, L., 2003. Development of nondestructive screening methods for single kernel characterization of wheat. Cereal Chemistry 80(3), 274-280.

Pasikatan, M.C., Dowell, F.E., 2004. High-speed NIR segregation of high- and low-protein single wheat seeds. Cereal Chemistry 81(1), 145-150.

Pedersen, D.K., Martens, H., Nielsen, J.P., Engelsen, S.B., 2002. Near-Infrared Absorption and Scattering Separated by Extended Inverted Signal Correction (EISC): Analysis of Near-Infrared Transmittance Spectra of Single Wheat Seeds. Applied Spectroscopy 56(9), 1206-1214.

Peltonen, J., Salovaara, H., 1991. Experiences of An Automatic Small-Scale Home Bakery in Test Bakings of 6 Spring Wheat-Varieties. Journal of Agricultural Science in Finland 63(2), 131-135.

Rittiron, R., Saranwong, S., Kawano, S., 2004. Useful tips for constructing a near infrared-based quality sorting system for single brown-rice kernels. Journal of Near Infrared Spectroscopy 12(2), 133-139.

Svenstrup, G., Heimdal, H., Nørgaard, L., 2005. Rapid instrumental methods and chemometrics for the determination of pre-crystallisation in chocolate. International Journal of Food Science and Technology 40, 953-962.

Triboi, E., Triboi-Blondel, A.M., 2001. Environmental effects on wheat grain growth and composition. Aspects Applied Biology 64, 91-101.

Tønning, E., Nørgaard, L., Engelsen, S.B., Pedersen, L., Esbensen, K.H., 2006. Protein heterogeneity in wheat lots using single-seed NIT - A theory of sampling (TOS) breakdown of all sampling and analytical errors. Chemometrics and Intelligent Laboratory Systems 84, 142-152.

Yoon, B.S., Brorsen, B.W., Lyford, C.P., 2002. Value of increasing kernel uniformity. Journal of Agricultural and Resource Economics 27(2), 481-494.

Zwingelberg, H., Brümmer, J.M., 1990. Backautomaten in Mühlenlaboratorien. Getreide, Mehl und Brot 44, 142-147.

Stepwise multivariate prediction of wheat flour functionality and bread quality

E. Tønning,[1,2,3] A. K. Thybo,[1] L. Pedersen,[4] F. van den Berg,[3]
S. B. Engelsen,[3] and L. Nørgaard[3]

[1] Plant Food Science, Department of Food Science, Faculty of Agricultural Sciences, University of Aarhus, DK-5792 Aarslev, Denmark, http://www.agrsci.org
[2] Corresponding Author. Phone: +45 8999-3413. Fax: +45 8999-3495. E-mail: erik.tonning@agrsci.dk
[3] Quality & Technology, Department of Food Science, Faculty of Life Sciences, University of Copenhagen, DK-1958 Frederiksberg C, Denmark, http://www.models.life.ku.dk
[4] Department of Chemical Engineering, Faculty of Engineering, University of Southern Denmark, DK-5230 Odense, Denmark, http://www.sdu.dk

ABSTRACT

Five spectroscopic technologies and combinations thereof were evaluated for prediction of twelve standard wheat flour parameters using partial least squares (PLS) and multi-block PLS (MBPLS). This was repeated on twenty-one quality attributes of breads baked from the flours to evaluate the feasibility for predicting end product quality at an early stage. Thirty-two diverse wheat lots were evaluated throughout an entire baking process from bulk grains to white bread. Near-infrared transmission (NIT) spectra were obtained from grains and flours. Near-infrared reflectance (NIR) and infrared (IR) spectra were obtained from the flours. The baking process was evaluated by nuclear magnetic resonance relaxometry of dough baked inside a temperature controlled magnet (NMR-baking). Flour functionality was assessed by protein and gluten contents, sedimentation, falling number, amylograph and farinograph recordings. Bread quality was assessed by mass, volume and instrumental and sensory texture analysis. NIT and NIR proved most powerful for prediction of flour quality. The flour quality parameters were indispensable for prediction of bread quality and when combined stepwise with NIT and NMR-baking data blocks the explained variance was further improved. The diversification of the wheat material and the baking method used were important for interpretation and require much more attention in the future.

Turning wheat into delicate and satisfactory spongy aromatic bread is an almost magic experience. The formation of the perfect visco-elastic dough capable of forming an air holding stable network of mainly protein and starch involves numerous chemical, physical and biological events to take place in timely order and intensity. Although bread making involves mixture of several ingredients, such as flour, yeast, water, oil, salt and sugar, it is the natural quantities and qualities of protein, starch and fat in the wheat which are the prerequisite for a good and dedicated bread product. The content and quality of the gluten fraction of the wheat protein is acknowledged to be the most important for the determination of end product quality of baked products (Bushuk et al 1969, Veraverbeke and Delcour 2002, Wesley et al 2001).

In the food industry there is an increasing demand for thorough quality specifications with respect to chemical, physical and biological properties of flour products (Mirablés 2004) to ensure end product specifications. Since flour functionality determination is a time consuming and laborious task involving wet chemistry and physics, the simultaneous prediction of some or all essential flour functionality measures from rapid spectroscopic techniques is of great interest. Standard biochemical parameters such as protein, wet gluten, moisture, ash and starch content are already routinely determined by near-infrared transmission spectroscopy with great accuracy. The prediction of protein composition from NIR is showing promising results. Wesley et al (2001) established calibrations for the glutenin and gliadin fractions and Fontaine et al (2002) presented good calibration with respect to individual amino acids. The feasibility of predicting rheological properties as determined by Farinograph, Mixograph and Alveograph measurements of dough using NIR has been assessed multiple times with varying success (Delwiche and Weaver 1994, Delwiche et al 1998, Dowell et al 2006, Hrušková et al 2004, Mirablés 2003

2004). Similarly the prediction of bread quality in terms of loaf volume, density and crump structure using NIR has been assessed (Delwiche and Weaver 1994, Delwiche et al 1998, Dowell et al 2006). The feasibility of establishing good prediction models with regard to the physical rheological properties is governed by the chemical information in the NIR region regarding protein content and composition (Wesley et al 2001). The success of those studies is apparently controlled by the diversity of the materials used. The most successful studies used very diverse materials (Delwiche et al 1998, Mirablés 2003, 2004) and it is noted that the ability to predict rheological properties is largely governed by the high correlation to crude protein content. Prediction of instrumental texture in bread has been assessed by nuclear magnetic resonance relaxation (Engelsen et al 2001).

Vibrational spectroscopy has proven extremely useful in the food and pharmaceutical industries for monitoring raw materials, process streams as well as end product quality (Dyrby et al 2002, Zachariassen et al 2005). Infrared as well as near-infrared spectroscopy provides a complex fingerprint of the internal chemical and physical composition of the irradiated food item which can be used for internal process monitoring and control in most applications. Vibrational spectroscopy measures the different vibrational energy states of molecular bonds. In the infrared region primarily the fundamental vibrational states are observed, while in near-infrared region it is the overtones and combination tones that are observed - primarily of anharmonic bonds that is C-H, N-H and O-H in biological materials. Although prediction of reference quality parameters is useful, the spectra contain much more information as they are complete maps of the material and in the case of living material a physicochemical map or fingerprint of the entire phenome (Munck 2005).

Another useful non-destructive technique is becoming more and more useful for qualifying food materials namely nuclear magnetic resonance (NMR) spectroscopy which studies the nuclear spin energy states of molecular nuclei, primarily the hydrogen (H) (Callaghan 1991). As the energy states of these nuclei are determined by the physical and chemical environment in which they are situated, very detailed information is available of complex systems such as food and natural products. One NMR technique is relaxometry in which H-atoms are exited in a magnetic field and the protons relaxation back to equilibrium is observed over time. The so called transverse relaxation (T_2-relaxation) can be observed by Carr-Purcell-Meiboom-Gill (CPMG) pulse train (Carr and Purcell 1954, Meiboom and Gill 1958) and has been used to observe water and fat compartmentalisation in dough and bread as well as throughout the baking process (Engelsen et al 2001) using a temperature ramp. By observing the fat and water dynamics taking place during the baking process more detailed information may be obtained of both the process as well as the flour quality. Thus NMR pose an interesting screening tool for dynamic fingerprinting of food stuffs during processing and cooking (Micklander et al 2002, Mortensen et al 2005) as well as baking and texture prediction (Engelsen et al 2001).

Normally standard baking tests require a fully equipped test bakery with trained staff for performing reproducible tests. In this study the use of automatic home-bakeries was adopted due to the overall good reproducibility (Hansen and Hansen 1992, 1993, Peltonen and Salovaara 1991) and their general use in for screening purposes in breeding test laboratories. Some report that the home-bakeries may lack discriminative power due to an unfortunate preference for weak doughs, which may mask the true potentials of wheats tested this way (Grausgruber et al 2001, Zwingelberg and Brümmer 1990). By ensuring large

diversity in the wheat materials analysed here (Tønning et al submitted), this was *a priori* considered a minor problem.

In process analytical technology (PAT) (USFDA 2004), a concept initiated in the pharmaceutical industry and now entering other industries including the food industry, multivariate data handling and analysis plays a central role due to the need for process understanding and control (Kourti 2006). New methods handling large and diverse data sets from several spectroscopic sources and process steps are needed and different multi-block solutions have been suggested. In classical calibration problems there are just two blocks of data, the descriptor **X**-block and the reference **Y**-block to be predicted. Relations between samples and variables within the **X** and **Y** blocks, respectively, can thus be evaluated by hierarchical modelling, such as principal component analysis (PCA) (Hotelling 1933, Wold et al 1987). The relation and calibration between **X** and **Y** can be handled by partial least squares regression (PLS) (Geladi and Kowalski 1986, Wold et al 1983). In modern process studies however, having just one **X** block and one **Y** block is unusual. Thus the need for more advanced modelling tools taking the structure and information of conceptually meaningful blocks into account is evident. An early application of multi-block PLS (MBPLS) concerning the combination of data from various spectroscopic sources for prediction of product quality dates back to 1984 (Frank et al 1984). MBPLS is basically a large PLS model in which data from conceptually meaningful blocks are treated separately with respect to pre-treatment, scaling and weighting relative to each other (Qin et al 2001, Westerhuis and Coenegracht 1997, Westerhuis et al 1998, Westerhuis and Smilde 2001). Other multi-block approaches exist in which common and unique information is separated for detailed assessment, Serial PLS (S-PLS) (Berglund and Wold 1999), Generalised orthogonal multiple co-inertia analysis PLS (GOMCIA-PLS)

(Vivien and Sabatier 2003), L-PLS (Martens et al 2005). The multiblock methods are generally applied to process monitoring (Choi and Lee 2005, Qin et al 2001, Westerhuis and Coenegracht 1997). Brás et al (2005) used MBPLS and S-PLS to compare and combine near-infrared and mid-infrared spectra for prediction of protein and moisture in soy beans. Felício et al (2005) performed similar near-infrared and mid-infrared experiments for gasoline and gas oil parameters, comparing single PLS, MBPLS and S-PLS. Vivien et al (2005) used GOMCIA-PLS in prediction of sensory data of peas from near-infrared data.

The aim of this study was to predict wheat flour functionality in terms of standard physicochemical parameters and wheat bread quality in terms of bread dimensions, instrumental texture and sensory evaluation by NIT spectroscopy of grain, NIT, NIR and IR spectroscopy of flour, NMR relaxation profiles of dough to bread and from milling process parameters. By using MBPLS, the individual and combined techniques are compared and the potential of complimentary information in different blocks is utilised (Fig. 1). In addition the sources of wheat material diversity are discussed in relation to the possibility of predicting quality parameters in flour as well as in bread.

MATERIALS AND METHODS

Wheat Material

Twenty-nine organic wheat lots from Aarslev (Aa) in Denmark and Kiel (Ki) in Germany grown in 2003 (03) and 2004 (04) were used. Two spring cultivars; Vinjett (Vi) grown in Aarslev and Combi (Co) grown in Kiel were diversified in the field by varying preceding catch crops. The catch crops were: No catch crop (NC), Winter Vetch (*Vicia villosa*) (WV), Fodder Radish (*Raphanus sativus var. Oleiformis*) (FR) and Turnip (*Brassica rapa*) (Tu). The last two are both from

the *Brassicaceae* family and were used interchangeably. Mixture lots from each year and location were formed and subsequently diversified in a laboratory scale TriQ single-kernel near-infrared (SKNIR) sorting device (BoMill AB, Lund, Sweden) according an internal baking quality trait into three quality fractions (F1, F2 and F3) (Löfqvist and Nielsen 2003). Two winter (Wi) cultivars; Capo, available from Kiel 2004 and Pentium, available from Aarslev 2004 was used without further diversification. Detailed information regarding the above mentioned twenty-nine wheat lots is available from Tønning et al (submitted). Additionally three conventionally grown lots from Sejet Planteforædling (Denmark) (Se) were included for reference. These lots were experimental cultivars assigned as feed (Feed) and bread wheats (Bre1 and Bre2) according to prior knowledge. The abbreviations in brackets above were used to form lot names for easy identification, thus a Vinjett lot from Aarslev 2004 grown after Fodder Radish has ID: 04AaViFR. The lots are presented in Table I with flour extraction rate (yield), ash content, protein content and falling number for reference.

Milling

The lots were milled on a laboratory scale mill, Brabander Quadromat Senior (Duisburg, Germany) separating the bran and germ from the endosperm. The process involved a conditioning step in which 0.4 to 2.2%-points of water was added depending on the actual moisture content of the grains. Moisture was determined by near-infrared transmission – see below. Grains and water were mixed in a rotating drum for 2.5 to 8.5 hours in order to soften the bran prior to the milling. The white flour particle size distributions were determined by controlled shaking of 100 g of flour through of a stack of sieves with consecutively smaller sieve sizes with two rubber cubes in each sieve for 10 min. Sieve seizes were 1000, 500, 250, 160 and 63 µm respectively. A total of

ten parameters were registered regarding the milling process ($\mathbf{X}_6$): Moisture content of the grain lots (12.1 – 15.4%), target conditioning moisture content (14.3 – 15.8%), conditioning water added (0.4 – 2.2%-points), conditioning time (2.5 – 8.5 h), relative milling yield of coarse bran (21.7 – 27.7%), fine bran (5.0 – 10.1%) and white flour (65.0 – 72.3%) and flour particle size <63 µm sieve size (17.7 – 31.7%), 63-160 µm sieve size (67.7 – 81.8%) and >160 µm sieve size (0.2 – 0.7%). The flour lots were kept in airtight plastic containers.

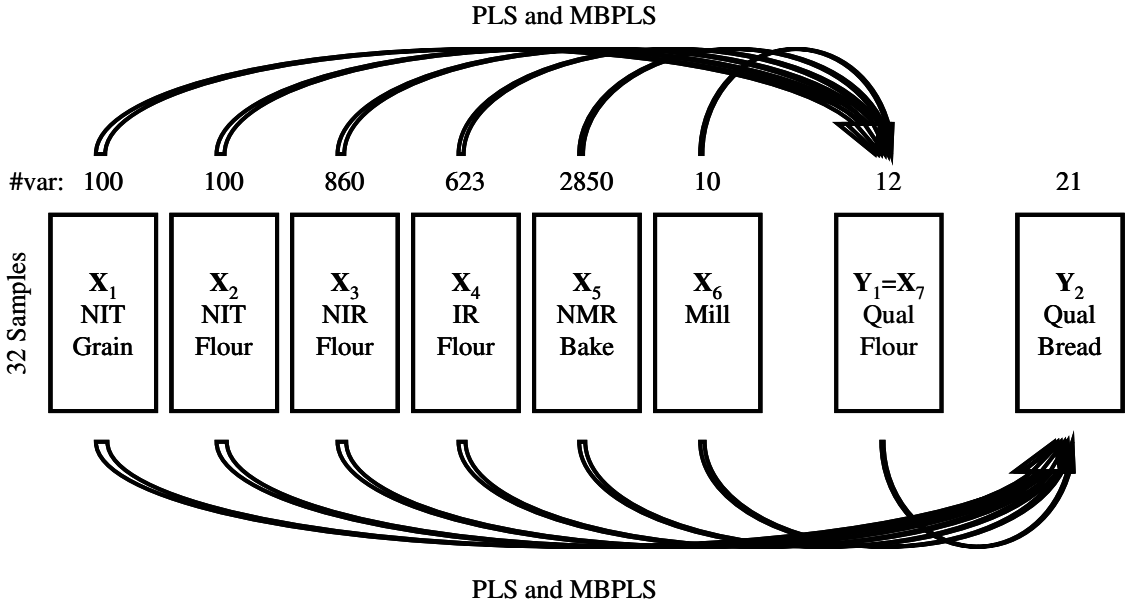


Fig. 1. Conceptual overview of the stepwise partial least squares (PLS) and multi-block (MBPLS) screening approach. The number of variables in each block is indicated above. Arrows indicate individual PLS models as well as MBPLS combing two or more blocks.

Grain NIT analysis

Approximately 800 g of grain representative for the lot was poured into a Foss Infratec 1241 Grain Analyser. Ten NIT spectra from 850 - 1048 nm for every 2 nm were recorded and averaged ($\mathbf{X}_1$). Bulk grain density (De_G) was automatically determined (Table II). Automatically predicted grain quality parameters were reported in Tønning et al (submitted).

TABLE I

Thirty-two wheat lots of six grouped according to origin and diversification. Arrows indicate how the mixture lot (Mx) in each group was sorted into approximately equally sized fractions by SKNIR. Flour yield, ash content, dry matter protein and falling number is given for reference. Further data may be found in Tønning et al (submitted).

Lot	M_{mix} [kg]	Sorting	M_{res} [kg]	Fl. Yield [%]	Ash DM [%]	Protein [%]	Fall. Num. [s]
2003 Aarslev Vinjett (03AaVi)							
No Catch (NC)	3.6 (20.8%)		4.1	71.3	0.53	12.0	376
Fodder Radish (FR)	6.0 (34.7%)		6.4	68.2	0.51	10.1	399
Winter vetch (WV)	7.7 (44.5%)		8.2	67.6	0.50	10.4	397
Mixture (Mx)	17.3 (100%)	→	3.8	71.8	0.61	10.6	365
Fraction 1 (F1)		→	4.1 (30.4%)	68.3	0.49	9.5	372
Fraction 2 (F2)		→	4.5 (33.3%)	68.1	0.50	10.4	404
Fraction 3 (F3)		→	4.9 (36.3%)	68.2	0.52	11.9	404
2004 Aarslev Vinjett (04AaVi)							
No Catch (NC)	6.7 (29.3%)		7.1	65.9	0.51	11.0	252
Turnips (Tu)	8.9 (38.9%)		9.4	66.6	0.51	9.6	234
Winter vetch (WV)	7.3 (31.9%)		7.8	69.2	0.53	10.1	234
Mixture (Mx)	22.9 (100%)	→	5.4	66.8	0.58	10.5	264
Fraction 1 (F1)		→	5.9 (35.1%)	69.0	0.50	8.8	212
Fraction 2 (F2)		→	5.5 (32.8%)	68.9	0.50	10.0	237
Fraction 3 (F3)		→	5.4 (32.1%)	69.4	0.51	11.9	250
2003 Kiel Combi (03KiCo)							
No Catch (NC)	5.3 (32.7%)		5.6	68.9	0.50	11.0	343
Fodder Radish (FR)	5.4 (33.3%)		5.9	68.1	0.49	10.5	423
Winter vetch (WV)	5.5 (34.0%)		5.8	66.1	0.49	10.7	395
Mixture (Mx)	16.2 (100%)	→	0.5*	n.a.	n.a.	n.a.	n.a.
Fraction 1 (F1)		→	4.9 (31.6%)	68.8	0.50	10.0	329
Fraction 2 (F2)		→	5.5 (35.5%)	69.3	0.51	11.2	408
Fraction 3 (F3)		→	5.1 (32.9%)	71.9	0.52	11.2	442
2004 Kiel Combi (04KiCo)							
No Catch (NC)	7.1 (31.0%)		7.1	66.5	0.48	9.6	272
Fodder Radish (FR)	8.4 (36.7%)		8.3	66.1	0.49	10.1	228
Winter vetch (WV)	7.4 (32.3%)		7.4	66.3	0.49	10.0	242
Mixture (Mx)	22.9 (100%)	→	4.7	68.6	0.52	10.1	251
Fraction 1 (F1)		→	5.3 (30.8%)	72.3	0.51	9.5	221
Fraction 2 (F2)		→	6.3 (36.6%)	70.3	0.49	9.8	241
Fraction 3 (F3)		→	5.6 (32.6%)	69.3	0.51	10.8	263
2004 Winter Crops							
04AaPent			N.a.	65.4	0.57	9.8	267
04KiCapo			N.a.	67.3	0.47	9.4	318
2004 Sejet conventional							
04SeFeed			N.a.	65.0	0.46	8.5	356
04SeBre1			N.a.	69.5	0.49	12.4	388
04SeBre2			N.a.	65.2	0.47	11.1	392

* This lot was too small and not included in the analysis

Flour NIT analysis

Two flour cups were filled for each lot with flour and inserted into the hopper of the Foss Infratec 1241 Grain Analyser. For each cup ten near-infrared transmission scans from 850 nm to 1048 nm were recorded and averaged ($\mathbf{X}_2$). Dry matter protein contents (Pr~F) along with other parameters reported elsewhere (Tønning et al submitted) were automatically calculated based on the Foss world-wide calibration.

Flour NIR analysis

Near-infrared reflectance spectra from 780 nm to 2498 nm in 2nm steps were recorded of flour filled ring cups on a Foss NIR Systems 6500 spectrometer with spinning sample module by averaging 16 scans. Background recordings were made using 8 scans. All lots were sampled in duplicates and subsequently averaged ($\mathbf{X}_3$).

Flour IR analysis

Infrared spectra from 1900 to 700 cm^{-1} for approximately every 2 cm^{-1} were recorded on an ABB/Bomen instrument with a diamond attenuated total reflection (ATR) unit. A small amount of flour was placed on the diamond and a pressure of 5 N/cm^2 was applied. A total of 64 scans were averaged by the instrument. Background recordings were made using 128 scans. All lots were sampled in duplicates and subsequently averaged ($\mathbf{X}_4$).

Standard flour test

The falling number (FN_F), was determined in duplicates using Standard No. 107/1 (ICC 1998) and averaged. Farinograph water absorption (WA_F), development time (DT_F), stability (St_F) and softening (So_F) were determined in a 50 g mixing chamber using Standard No. 115/1 (ICC 1998).

Zeleny sedimentation volume (Ze_F) was determined in duplicates according to Approved Method 56-61A (AACC 2000) and Standard No. 116/1 (ICC 1998) and averaged. Amylograph gelatinisation maximum (GM_F) and temperature (GT_F) at maximum gelatinisation was determined in duplicates according to Standard no. 126/1 (ICC 1998) and Approved Method 22-10 (AACC 2000) and averaged. Wet gluten (WG_F) and gluten index (GI_F) were determined in duplicates according to Approved Method 38-12 (AACC 2000) and Standard No. 155 (ICC 1998) and averaged (Table II). The data was stored in data matrix $\mathbf{Y}_1=\mathbf{X}_7$ and with bulk grain density (De_G) and protein content (Pr~F).

Dough preparations and bread baking

For each lot, three breads were made in household baking machines (Grausguber et al 2001, Hansen and Hansen 1992, 1993, Zwingelberg and Brümmer 1990), Panasonic Automatic Bread Bakery, SD-253 in the following manner. 7.1 g of dry yeast was distributed evenly in the pan around the kneading blade. Then 600 g of flour on a 14% moisture base were added followed by 8.4 g table sugar and 8.4 g salt. 5 mL 0.48% ascorbic acid solution corresponding to 40 mg ascorbic acid/kg flour and tap water at 25ºC up to total water amount corresponding to Farinograph water absorption was added. The baking process was initiated using the following SD-253 program settings, Size: XL, Bread type selection: BASIC, Option: BAKE RAPID and Crust colour: MEDIUM. The basic program ran for 115 min. as follows: 20 min. kneading 20 min. rising, 1 min. kneading, 39 min. rising and 35 min. of baking. Immediately after second kneading at 40 min, approximately 2 g of dough was transferred from one of the baking machines to a 13 mm vide glass vial which was inserted into an 18 mm sized glass NMR tube and closed with a plastic cap and placed in the variable NMR temperature probe for NMR-baking (see below).

Immediately after baking, the breads were removed from the baking pans, weighed and left on a wire rack to cool for 90 minutes. The breads were weighed again (MaFB) and the longitudinal as well as the transverse circumference, c_L and c_T, over the bread centre were measured. The volume of the bread was estimated by assuming a cuboid shaped bread with length and width determined by the pan size, 175 mm × 130 mm, and the height determined from circumferences:

$$\text{Height} = \frac{c_L/2 - 175\text{mm} + c_T/2 - 130\text{mm}}{2} \qquad \text{(Eq. 1)}$$

This cuboid assumption does not take the individual more or less rounded shapes with cracks and pores of the bread crusts into account, but serves as an adequate substitute for the actual volume (VoFB) and for density (DeFB) calculation. Additionally the relative water loss during baking and cooling (WLFB) was registered (Table II). The data was stored in matrix $\mathbf{Y}_2$ along with instrumental texture analysis and sensory analysis. Two breads were frozen to -18ºC for sensory analysis of thawed bread (See below), while one was used immediately for texture analysis of the fresh bread crumb (See below).

Dynamic NMR-baking profiling

A 2 g dough sample (above) was rised and baked in a 23.2 MHz Maran Benchtop pulsed NMR Analyser, Resonance Instruments (Whitney, UK). CPMG recordings were made using the following settings: dwell time, DT = 50 µs, echo time, τ = 100 µs, number of echoes, NE = 8190, points recorded for every echo, NECH = 1, receiver delay, RD = 1s, number of scans, NS = 8. Only even numbered echoes were used. Instrument dead time was 10 µs. The temperature was controlled by a continuous flow of air in the range from 34ºC to 96ºC and back to 34ºC mimicking the temperature at the centre of the bread

during rising, baking and cooling. A measurement were made every 4 min for 120 minutes, thus a total of 30 CPMG curves were recorded. The temperature was regulated after each measurement allowing approximately 3 min for stabilising the internal sample temperature. Ten recordings were made at 34ºC (raising time), then one recording at each of the temperatures, 36, 41, 46, 51, 56, 61, 66, 71, 76, 81, 86, 91 and 96ºC (baking) and finally seven recordings at 34ºC (cooling). Before multi-block analysis the redundant data were reduced from 4095 time points to 95 time points for each relaxation curve in the following way. The initial time points 1 to 20 was were kept without alterations, time points 21 to 95 were reduced to 25 points by averaging every three points, time points 96 to 495 were reduced to 25 point by averaging every 16 points, and time points 496 to 4095 were reduced to 25 points by averaging every 144 points. Every NMR baking profile (30x95 data points) was max-normalised to the first acquisition point of the first CPMG curve to even out possible sample mass variations in the measurement area of the magnet. Relaxation curves were concatenated prior to the multi-block analysis ($\mathbf{X}_5$) (Fig. 2E).

Texture profile analysis of fresh bread crumb

From each lot, one bread was cut into nine approximately 20 mm slices on a slicing machine. Actual slice thicknesses varied between 17.9 and 23 mm. Only slice no. 2, 3, 7 and 8 were used for texture profile analysis (TPA). The end slices 1 and 9 as well as centre slices 4, 5 and 6 were discarded to avoid influences from the crust and the kneading blade perforation respectively. The TPA was conducted by placing an entire slice of bread on the base plate of the TA-XT2 Texture Analyser, Stable Micro Systems fitted with a cylindrical 40 mm SMS steel probe with sharp edges. The force used to compress and decompress the bread crumb 15 mm (75%) twice with a 5 s delay between the first and the second compression was recorded. Test speed was 1 mm s^{-1}, trigger

force 0.02 N and post test speed was 5 mm s^{-1} after second down stroke. From each time-force curve and actual slice thickness from the TA-XT2 the following nine parameters were extracted and calculated and the results from the four slices were averaged:

Hardness 1 at 60% compression (H1): Force at 60 % compression during first down stroke corrected for deviation from 20 mm slice thickness. H1 = $F_{60\%,20\text{mm}}$. The force at 60% compression, $F_{60\%}$, and thickness, d in mm, were given by the TA-XT2. Thus the force expected to be used on the individual slices if they would have been 20 mm, $F_{60\%,20\text{mm}}$, can be calculated:

$$F_{60\%,20\text{mm}} = F_{60\%} \frac{20\text{mm}}{d} \qquad \text{(Eq. 2)}$$

Thus a linear relationship is expected between thickness and force used to compress a given sample 60%. This assumption holds as long as deviations in thickness are small and improve the relative standard deviation on the hardness measure. Hardness 2 at 60% compression (H2): Determined in the same way as H1 above, only for second down stroke. The 60% compression corrected with sample thickness was used as a more robust hardness measures in stead of the less robust maximum compression (75%) which due to the high compression rate is very sensitive to deviations in sample height. Adhesive Force (AF): The maximum negative force exhibited during first up stroke. Adhesiveness (Ad): The work done by the bread crumb to hold on to load cell, i.e. the negative area under the curve, during first up stroke. Resilience (Re): Area under curve during first upstroke until zero force divided by area under curve during first down stroke. Cohesiveness (Co): Area under curve during second down stroke divided by area under curve during first down stroke. Springiness (Sp): Distance travelled during second down stroke divided by distance travelled under first

compression (15 mm). Gumminess (Gu): Gu = Co·H1. Chewiness (Ch): Hh = Gu·Sp (Table II). The data was stored in matrix $\mathbf{Y}_2$.

Sensory evaluation of thawed bread crumb

One bread was thawed at room temperature and equilibrated for 16 hours before cut into 20 mm slices. Sensory profiling was performed in a sensory evaluation laboratory according to international standards (ASTM STP 913). The basic trained panel with many years of sensory experience was composed of 10 assessors (4 males/6 females, aged from 30-55 years). They were trained for 6 hours in sensory profiling of bread. Bread slices were cut vertically in halves and served to the assessors. Thirty-two bread samples were served in 5 sensory sessions with 6 or 7 samples, respectively. The panel evaluated 10 sensory attributes: yellowness, elasticity, fracturability, hardness, dryness, chewiness, adhesiveness and after taste, which are described in Table II. Two attributes, wheat bread aroma and wheat bread taste intensities were discarded, as no systematic variation was found. Results from one assessor were likewise discarded due to low signal to noise ratio on all attributes. Sensory analysis was repeated another day on the second bread. Replicate variation as well as assessor variation in the use of scale was removed by ANOVA-PLS (A-PLS) with **X** as replicates and assessors 1 to 10 assigned with either zeroes or ones and **Y** as sensory data. The **Y** residual after 11 PCs was extracted from the A-PLS result and was hereafter used as level corrected sensory data in the subsequent analysis (Martens et al 2000) (Table II). The data was stored in matrix $\mathbf{Y}_2$.

TABLE II

Quality parameters for flour functionality (Y_1) and bread quality (Y_2).

Block	Parameter	Code	Description	Min	Max	Mean	Std
Y_1	**Standard flour and grain tests ($Y_1 = X_7$)**						
	Density	De_G	Density of bulk grains determined by weight module [kg/hL]	71.2	86.1	79.5	3.9
	Protein (DM)	Pr-F	Dry matter protein concentration in flour determined by NIT, Foss Infratec 1241 [%]	8.5	12.4	10.4	0.9
	Falling Number	FN_F	Falling number determined at 14% moisture [s]	212	442	318	75
	Water Absorption	WA_F	Farinograph water absorption in flour determined at 14% moisture [%]	51.1	61.8	56.6	3.4
	Development Time	DT_F	Farinograph development time in flour determined at 14% moisture [min]	1.3	3.0	1.9	0.4
	Stability	St_F	Farinograph stability in flour determined at 14% moisture [min]	0.9	8.0	3.6	2.2
	Softening	So_F	Farinograph softening in flour determined at 14% moisture [FU]	45	155	85	26
	Zeleni	Ze_F	Sedimentation test at 14% moisture in flour [mL]	25.4	40.9	31.7	3.5
	Gelatinisation Temp.	GT_F	Amylograph gelatinisation temperature in the gel. maximum [ºC]	70.4	91.1	81.3	7.9
	Gelatinisation Max.	GM_F	Amylograph gelatinisation maximum [AU]	195	1630	705	494
	Wet Gluten	WG_F	Wet gluten in grain determined at 14% moisture [%]	14.8	29.8	21.6	3.3
	Gluten Index	GI_F	Gluten quality index, percentage strong gluten [%]	71.2	100.0	95.4	7.3
Y_2	**Bread proportions**						
	Mass	Ma_B	Mass of bread after cool down [g]	855	923	887	21
	Relative Water Loss	WL_B	Relative water loss during baking and cool down [%]	19%	27%	22%	2%
	Volume	Vo_B	Bread volume estimated from circumference after cool down [L]	2.47	3.87	3.14	0.42
	Density	De_B	Based om volume and mass after cool down [g/L]	0.23	0.37	0.29	0.04
	Texture profil analysis on fresh bread using TA-XT2 Texture Analyser						
	Hardness 1	H1FB	Force recorded at 60% compression during first down stroke [N]	3.2	14.5	7.6	3.1
	Hardness 2	H2FB	Force recorded at 60% compression during second down stroke [N]	2.4	9.1	5.1	2.0
	Adhesive Force	AFFB	Maximum negative force recorded during first up stroke [N]	0.02	0.71	0.19	0.14
	Cohesiveness	CoFB	Work done during second down stroke relative to work done during first down stroke [-]	0.57	0.73	0.64	0.04
	Springiness	SpFB	Recovery height relative to compression length [-]	0.84	0.96	0.90	0.03
	Gumminess	GuFB	Gumminess = Cohesiveness × Hardness 1 [N]	2.2	8.4	4.8	1.8
	Chewiness	ChFB	Chewiness = Gumminess × Springiness [N]	2.0	7.5	4.3	1.5
	Resilience	ReFB	Positive work done during first upstroke rel. to work done during first down stroke [-]	0.21	0.36	0.27	0.03
	Adhesiveness	AdFB	Negative work done during first upstroke [Ns]	0.02	5.18	0.75	0.96
	Sensory panel evaluations on thawed bread corrected for replicate and assesor level effects						
	Yellowness	YeTB	Rating of the colour yellow in the crumb	-4.8	5.1	0.1	3.2
	Elasticity	ElTB	Recovery of crumb height rated after compression between thumb and forefinger	-3.6	3.5	0.0	2.1
	Fracturability	FrTB	The tendensy of the crumb to fracture while stretching the crumb between the hands	-3.6	4.1	0.1	2.0
	Hardness	HaTB	The hardness of crumb rated at first bite	-2.9	3.5	0.0	1.8
	Dryness	DrTB	Rating of the crumb dryness in the mouth when chewing	-3.1	2.8	-0.1	1.4
	Chewiness	CwTB	The amount of mastications before the crumb bite is ready to swallow	-2.2	2.8	0.0	1.2
	Adhesiveness	AhTB	The tendency of the crumb to stick in the mouth and forming a lump when chewing	-1.9	1.9	0.0	0.9
	Aftertaste	AfTB	Intensity of non-specific after-taste (e.g. sourish, staled, yeast)	-1.2	2.9	0.0	1.2

Data analysis

NIT spectra of grains and NIT, NIR and IR spectra of flours were pre-processed in the Unscrambler (Version 9.2, Camo, Norway) using extended multiplicative scatter correction (EMSC) in order to remove additive, multiplicative and channel and squared channel dependent scatter resulting from physical variations in the samples (Martens et al 2003, Martens and Stark 1991). The spectroscopic data ($\mathbf{X}_1$ – $\mathbf{X}_5$) were centred and the milling data ($\mathbf{X}_6$), flour functionality ($\mathbf{Y}_1$) and bread quality data ($\mathbf{Y}_2$) were centred and scaled to unit variance.

The multivariate data analysis was separated into two parts. In the first part MBPLS was used to screen the ability of the five spectroscopic techniques ($\mathbf{X}_1$ – $\mathbf{X}_5$) and the milling conditions ($\mathbf{X}_6$) individually and in combinations to predict twelve physicochemical flour parameters ($\mathbf{Y}_1$). In the second part, the twelve physicochemical variables were transferred to the explanatory blocks as $\mathbf{X}_7$. $\mathbf{X}_1$ – $\mathbf{X}_7$ was used to predict twenty-one bread quality attributes ($\mathbf{Y}_2$). $\mathbf{Y}_2$ contained four dimensions, nine TPA attributes and eight sensory perceived attributes. A conceptual overview of the block structure is presented in Fig. 1. The PLS and MBPLS regression models were calculated using the Multiblock-Toolbox (van den Berg et al 2001) from www.model.life.ku.dk in Matlab (Version 6.5.0.180913a Release 13, The MathWorks, Inc., USA) and validated with full cross validation. The predictor blocks ($\mathbf{X}_1$ – $\mathbf{X}_7$) were equally weighted to norm 1.

RESULTS AND DISCUSSIONS

In Fig. 2 the eight different data-blocks ($\mathbf{X}_1$-$\mathbf{X}_6$, $\mathbf{Y}_1$-$\mathbf{Y}_2$) for the 32 samples are presented after pre-processing and centring. NIT of grain and flour, NIR and IR

of flour ($\mathbf{X}_1$-$\mathbf{X}_4$) are presented after EMSC and centring, NMR-baking profiles ($\mathbf{X}_5$) are shown as concatenated and centred relaxation curves and milling process parameters ($\mathbf{X}_6$) as well as flour quality parameters ($\mathbf{Y}_1$=$\mathbf{X}_7$) and the bread quality parameters ($\mathbf{Y}_2$) are presented as scaled to unit variance and centred. An overview of the major variance in all blocks are shown in eight corresponding PCA scores plots in Fig. 3 showing the first two principal components. The sample origins are present as more or less distinct groups in all plots, however, the groups are not located relative to each other in the same way, which indicate that the individual blocks may contain both similar and complementary information regarding the lots. Within each group of spring wheats (03Ki, 03Aa, 04Ki and 04Aa) which were diversified post harvest using the TriQ SKNIR quality sorter the resulting fractions are connected with lines. Especially the flour functionality block ($\mathbf{Y}_1$) in Fig. 3G contains patterns influenced markedly by the quality diversification by fractionation. This systematic relation is not as pronounced in the remaining plots at least not in the first two PCs presented, although $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$ and $\mathbf{X}_4$ in Fig 3A, B, C and D do indicate some systematic variation with respect to the fractionation. It is noteworthy that the post harvest sorting fractions in Fig. 3G are spanning much more variation in each group than the starting material varied by agronomical measures using different preceding catch crops. Finding the flour functionality greatly influenced by the fractionation is very encouraging in terms of the potential for post harvest diversification for added value. This was investigated further in Tønning et al (submitted).

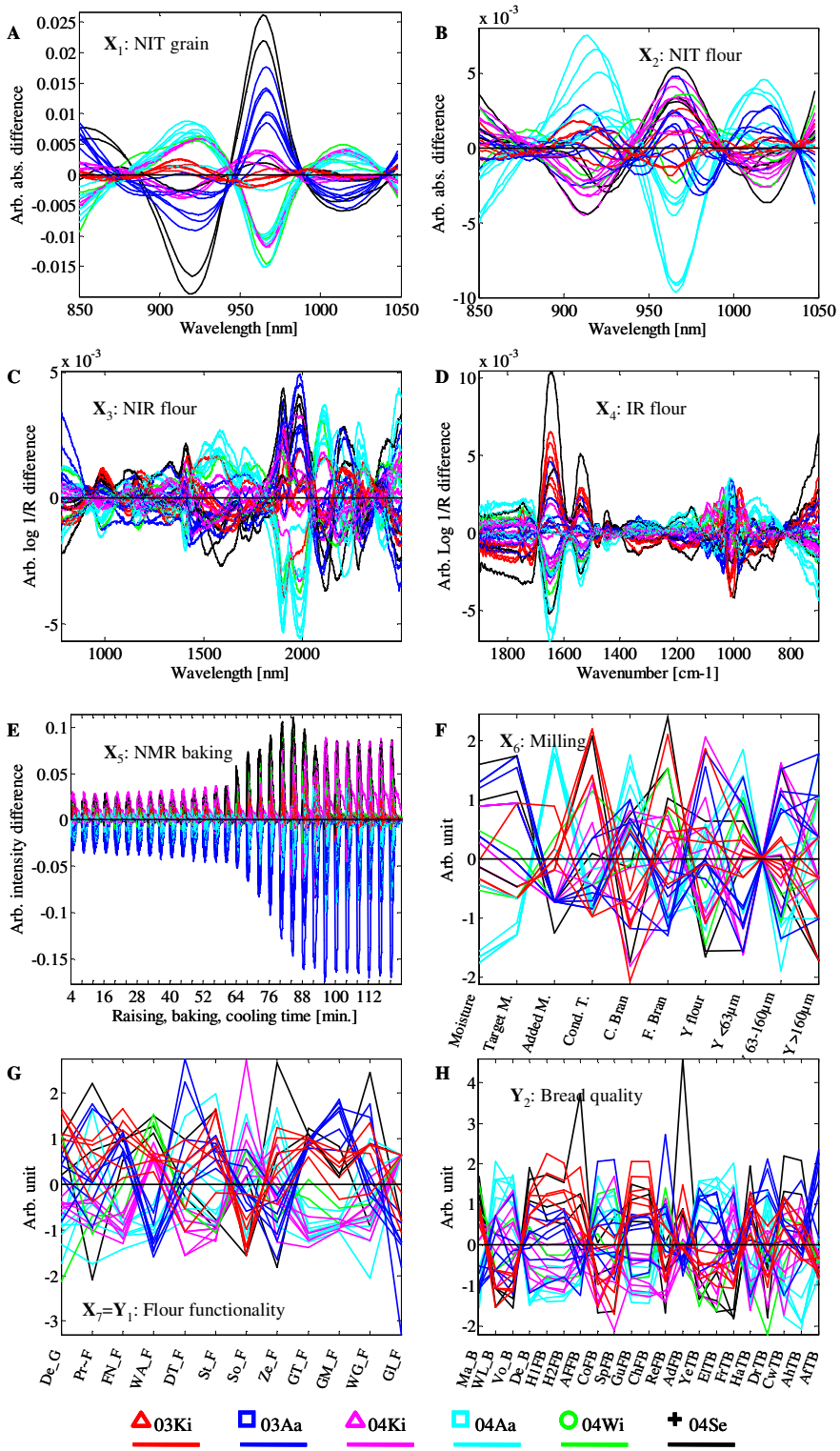


Fig. 2. Eight blocks of data. A: EMSC and centred NIT of grains, B: EMSC and centred NIT of flour, C: EMSC and centred NIR of flour, D: EMSC and centred IR of flour, E: Centred NMR-baking profiles of dough, F: Auto-scaled milling parameters, G: Auto-scaled flour quality parameters and H: Auto-scaled bread quality parameters. Colours according to origin (see text and legend in figure and parameter abbreviations in Table II).

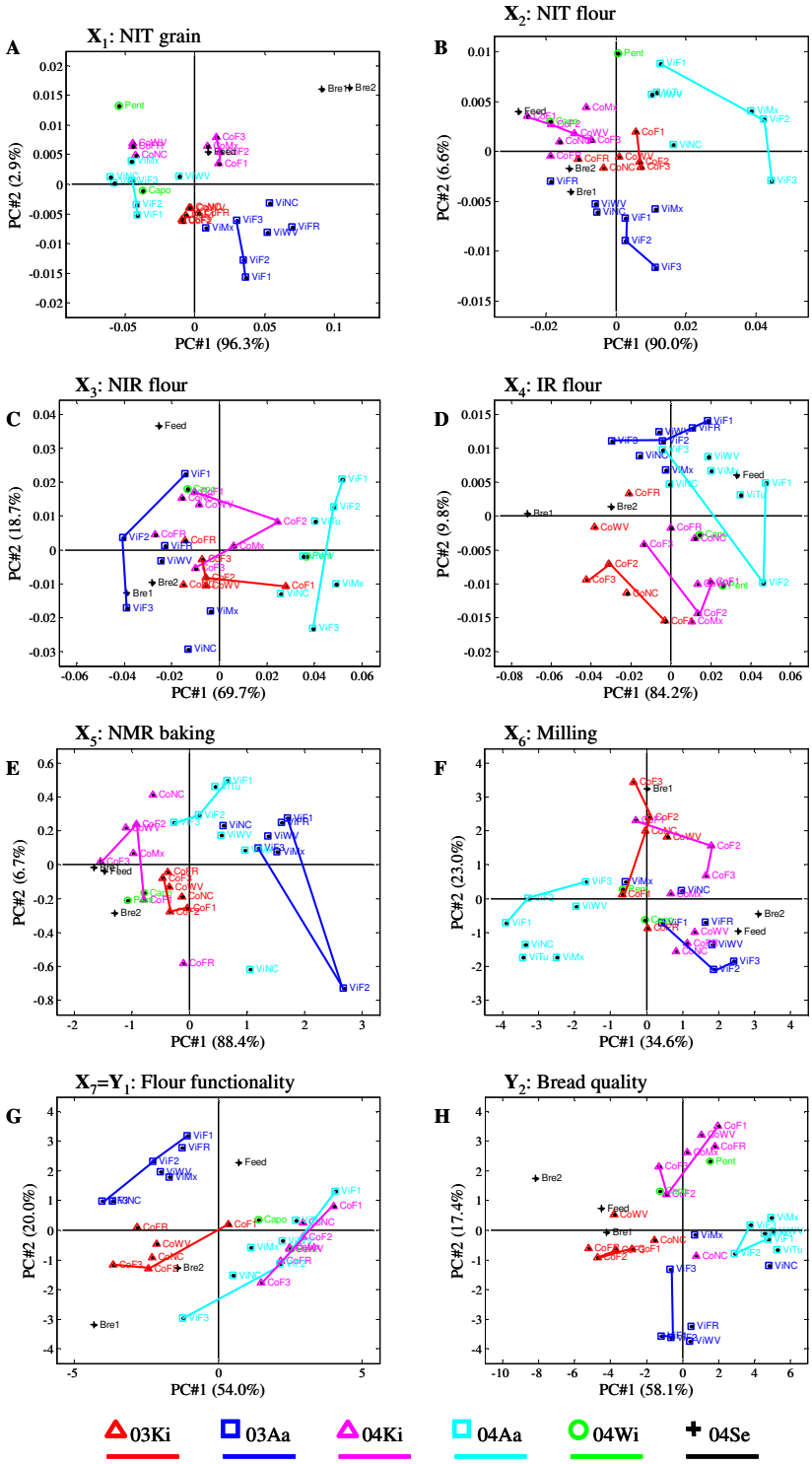


Fig. 3. Eight PCA plots of the data blocks presented in Fig. 1. A: EMSC and centred NIT of grains, B: EMSC and centred NIT of flour, C: EMSC and centred NIR of flour, D: EMSC and centred IR of flour, E: Centred NMR-baking profiles of dough, F: Auto-scaled milling parameters, G: Auto-scaled flour quality parameters and H: Auto-scaled bread quality parameters. Colours according to origin (see text and legend in figure and sample abbreviations in Table I).

The prediction of flour functionality parameters in the $\mathbf{Y}_1$-block were investigated using the spectroscopic methods and the milling parameters ($\mathbf{X}_1$-$\mathbf{X}_6$) as predictor blocks with PLS and MBPLS. In Fig. 4A the explained variance of $\mathbf{Y}_1$ can be viewed using separate **X**-blocks as well as the combination of NIT of grain and NIT of flour ($\mathbf{X}_1\mathbf{X}_2$). Other combinations of two and three blocks were also investigated, however not displayed, since combining other blocks did not improve the overall model performance. With single blocks, the best predictions were obtained using either NIT on grain ($\mathbf{X}_1$), NIT on flour ($\mathbf{X}_2$) or NIR on flour ($\mathbf{X}_3$). They model 70.2%, 73.6% and 78.9% of $\mathbf{Y}_1$ using six, nine and ten PLS components respectively. In comparison IR ($\mathbf{X}_4$) were performing less satisfactory explaining only 68% of $\mathbf{Y}_1$. This confirms the near-infrared advantages over the mid-infrared (MIR) for prediction of internal quality traits as also noted by Brás et al (2005) arguing that the MIR region has a lower penetration depth and thus is very susceptible to variations in particle size distribution and general sampling and sample presentation issues due to very small sample size. The NMR-baking profiles ($\mathbf{X}_5$) were generally unable to describe flour functionality. The milling process parameters ($\mathbf{X}_6$) had no effect on (correlation to) the flour functionality ($\mathbf{Y}_1$) parameters presented here, hence conveniently ruling out the milling process as a source for systematic variation in the flour quality.

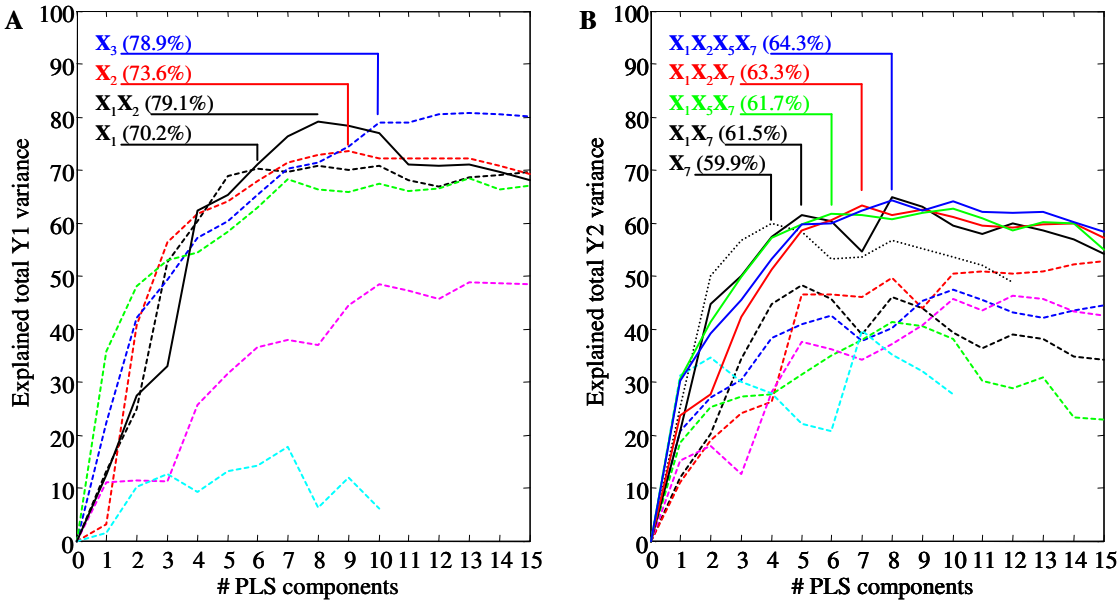


Fig. 4. PLS and MBPLS modelling performances. A: Explained flour functionality variance vs. number of PLS (MBPLS) components using individual predictor blocks and selected combinations. $\mathbf{X}_1$: ---, $\mathbf{X}_2$: ---, $\mathbf{X}_3$: ---, $\mathbf{X}_4$: ---, $\mathbf{X}_5$: ---, $\mathbf{X}_6$: ---, $\mathbf{X}_1\mathbf{X}_2$: —. B: Explained bread quality variance vs. number of PLS (MBPLS) components using individual predictor blocks and selected combinations. $\mathbf{X}_1$: ---, $\mathbf{X}_2$: ---, $\mathbf{X}_3$: ---, $\mathbf{X}_4$: ---, $\mathbf{X}_5$: ---, $\mathbf{X}_6$: ---, $\mathbf{X}_7$: ····, $\mathbf{X}_1\mathbf{X}_7$: —, $\mathbf{X}_1\mathbf{X}_2\mathbf{X}_7$: —, $\mathbf{X}_1\mathbf{X}_5\mathbf{X}_7$: —, $\mathbf{X}_1\mathbf{X}_2\mathbf{X}_5\mathbf{X}_7$: —.

By using MBPLS, it was possible to combine two or more blocks for prediction of $\mathbf{Y}_1$. The blocks were weighted equally to norm 1, regardless of the number of variables in the individual blocks. Thus NIT of grains containing 100 variables had the same weight in the modelling as the NMR-baking profiles with 2850 variables. Although different weighting could be used, equal weighting was a pragmatic solution when the relative importance of the individual blocks was unknown. If two blocks contain complimentary information about $\mathbf{Y}_1$ variables, the degree of explanation is expected to increase, when the blocks are combined, while combining block containing only similar information about $\mathbf{Y}_1$ a worse or an equal model is expected. By combining NIT of grain ($\mathbf{X}_1$) and NIT of flour ($\mathbf{X}_2$), the overall explained $\mathbf{Y}_1$-variance was increased to 79.1% using eight principal components (Fig. 4A). The NIT technique was used on grain and flour,

hence also conceptually complementary which is important in order to gain better model performance by combining blocks (van den Berg et al 2001). NIT on grain was probably particular important as some of the major variations in $\mathbf{Y}_1$ seen in Fig. 3G were due to fractionation based on single-kernel near-infrared technology on the raw material (Löfqvist and Nielsen 2003, Tønning et al., submitted). Other blocks may contain complimentary information, but while this was not related to $\mathbf{Y}_1$, combining those blocks with NIT or NIR blocks ($\mathbf{X}_1$-$\mathbf{X}_3$) did not improve the explained variance of $\mathbf{Y}_1$.

In Fig. 4B the accumulated explained variances of the twenty-nine bread quality ($\mathbf{Y}_2$) attributes are shown for the seven individual predictor blocks, now including flour quality parameters ($\mathbf{Y}_1$) as the seventh predictor block ($\mathbf{X}_7$). While $\mathbf{X}_1$ to $\mathbf{X}_6$ all explain less than 50% of the bread quality variance respectively, the flour quality block distinctively explains 59.9% of the variance regarding bread quality using only four PLS components. Although the spectroscopic techniques in principle are truly physicochemical fingerprints of the entire sample physics and chemistry, they are outperformed as predictors by the twelve standard flour tests. However by combining $\mathbf{X}_7$ with NIT on grain ($\mathbf{X}_1$), the prediction was improved to 61.5% using five components. Further addition of NMR-baking profiles ($\mathbf{X}_5$) or NIT of flour ($\mathbf{X}_2$) improved explained $\mathbf{Y}_2$ variance to 61.7% with six PLS components and 63.3% with seven PLS components respectively. Combining all above mentioned blocks ($\mathbf{X}_1\mathbf{X}_2\mathbf{X}_5\mathbf{X}_7$) the explained $\mathbf{Y}_2$ variance was increased to 64.3% with eight PLS components. Thus the conceptually different blocks do contain relevant complementary $\mathbf{Y}_2$-related information.

The rough evaluations in Fig. 4 of the various **X**-blocks and their combinations as predictors for the total **Y**-variances do not provide detailed information of the

ability for predicting individual quality parameters. In Table III, $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$ and $\mathbf{X}_1\mathbf{X}_2$ are evaluated with respect to model complexity, i.e. number of optimal PLS components, correlation coefficient (*r*), *RMSECV* and explained $\mathbf{Y}_1$-variations for the optimal models in each response variable.

TABLE III

Optimal number of PCs for individual flour functionality parameters using the four PLS and MBPLS models with the best overall performances (abbreviations in Table II).

	$\mathbf{X}_1$: NIT grain				$\mathbf{X}_2$: NIT flour				$\mathbf{X}_3$: NIR flour				$\mathbf{X}_1\mathbf{X}_2$: NIT grain + flour			
	#Comp	*r*	*RMSECV*	%Y	#Comp	*r*	*RMSECV*	%Y	#Comp	*r*	*RMSECV*	%Y	#Comp	*r*	*RMSECV*	%Y
De_G	5	0.83	2.2	68.6	7	0.90	1.74	80.0	6	0.92	1.51	84.9	**8**	**0.94**	**1.37**	**87.6**
Pr~F	11	0.96	0.26	91.6	**5**	**1.00**	**0.089**	**99.0**	6	0.99	0.090	99.0	10	0.99	0.102	98.7
FN_F	6	0.93	27	86.5	12	0.89	34	79.1	**10**	**0.94**	**24**	**89.1**	8	0.93	27	86.9
WA_F	13	0.95	1.02	90.8	10	0.96	0.91	92.7	**12**	**0.98**	**0.70**	**95.7**	13	0.98	0.70	95.6
DT_F	**4**	**0.80**	**0.23**	**63.8**	2	0.74	0.26	54.7	4	0.68	0.29	44.5	5	0.75	0.26	55.1
St_F	5	0.85	1.15	71.7	3	0.79	1.34	62.0	**12**	**0.86**	**1.14**	**72.5**	7	0.82	1.27	65.8
So_F	5	0.76	16.5	57.8	8	0.83	14.2	68.6	3	0.71	17.9	50.3	**8**	**0.88**	**12.1**	**77.2**
Ze_F	8	0.79	2.1	61.8	4	0.86	1.73	74.2	**7**	**0.89**	**1.58**	**78.6**	4	0.84	1.83	71.2
GT_F	7	0.94	2.6	88.9	13	0.89	3.6	78.5	**13**	**0.96**	**2.2**	**92.2**	9	0.94	2.6	89.0
GM_F	10	0.93	178	86.6	11	0.91	199	83.3	**12**	**0.97**	**111**	**94.8**	8	0.94	172	87.6
WG_F	10	0.97	0.85	93.3	10	0.96	0.87	92.8	**8**	**0.98**	**0.67**	**95.8**	8	0.97	0.75	94.8
GI_F	5	0.73	5.0	52.4	6	0.75	4.9	54.7	**8**	**0.77**	**4.7**	**57.2**	7	0.76	4.7	57.0

The optimal models are highlighted and the corresponding predicted vs. measured plots are presented in Fig. 5A-L. Clearly NIR of flour ($\mathbf{X}_3$, Table III) is the individual technique which provides the best modelling power as was concluded from Fig. 4A. Protein (Pr~F) and wet gluten (WG_F) contents are very well determined with 99.0% and 95.9% explained using relatively few PLS components, six and eight respectively. A clear diversification in protein and gluten contents was achieved by the TriQ fractionation (Fig. 5B and K). Amylase activity measures; falling number (FN_F), gelatinisation temperature and maximum (GT_F, GM_F), were also well predicted with 89.1%, 92.2% and 94.8% explained, however using ten, thirteen and twelve PLS components respectively. In comparison Dowell et al (2006) was unable to predict falling number using near-infrared reflection and transmission. From Fig. 5C, I and J it is clear that each group had a distinct level of amylase activity, high in 2004 and low in 2003, and although diversified further by the TriQ fractionation, the groups are clearly determining the outcome of the regression. Hence the good

correlation to amylase activity could at least partially be caused by confounding. The correlation may thus have and indirect origin or component not specifically connected to the amylase activity itself, but rather relate to other factors discriminating the groups. Farinograph water absorption (WA_F) was well predicted with 95.9% explained variance using twelve PLS components due to great differences with respect to the location and cultivar as well as TriQ fractionation effect (Fig. 5D). The associated Farinograph parameters, development time (DT_F), stability (St_F) and softening (So_F) were less well predicted although a clear effect of TriQ fractionation is observed in Fig. 5E, F and G. Softening was best predicted by combining NIT of grain and flour ($\mathbf{X}_1\mathbf{X}_2$) explaining 77.2%, while stability was explained by 72.5% using NIR on flour ($\mathbf{X}_3$) and development time was explained by 63.8% using NIT on grain ($\mathbf{X}_1$). In comparison Dowell et al (2006) and Mirablés (2004) got similar result for water absorption predicted by NIT however Mirablés (2004) provided much better predictions for the remaining Farinograph parameters, while Hrušková et al (2004) got worse result using NIR. Zeleny sedimentation was explained by 78.6% using NIR ($\mathbf{X}_3$) and had a marked effect of the TriQ fractionation (Fig. 5H). This was similar to Delwiche et al (1998) and Dowell et al (2006). The unfortunate distribution of gluten index (GI_F) with two thirds of the lots having GI=100% resulted in poor predictions (Fig. 5L). It is quite clear from Table III and Fig. 5, that the rough overall evaluation made in Fig. 4A suggesting an optimal model complexity using six, nine, ten and eight principal components was inadequate to determine the actual best predictor block or combination and the optimal model complexity for prediction of the individual parameters. Although a combination of NIT of grain ($\mathbf{X}_1$) and NIT of flour ($\mathbf{X}_2$) appeared superior, NIR on flour ($\mathbf{X}_3$) was performing best with regards to individual parameters. It is quite clear that the diversity of the lots included is very important for establishing good prediction models (Delwiche et al 1998),

however also the origin of the variation as presented in Fig. 5A-L may be important in order to gain process understanding.

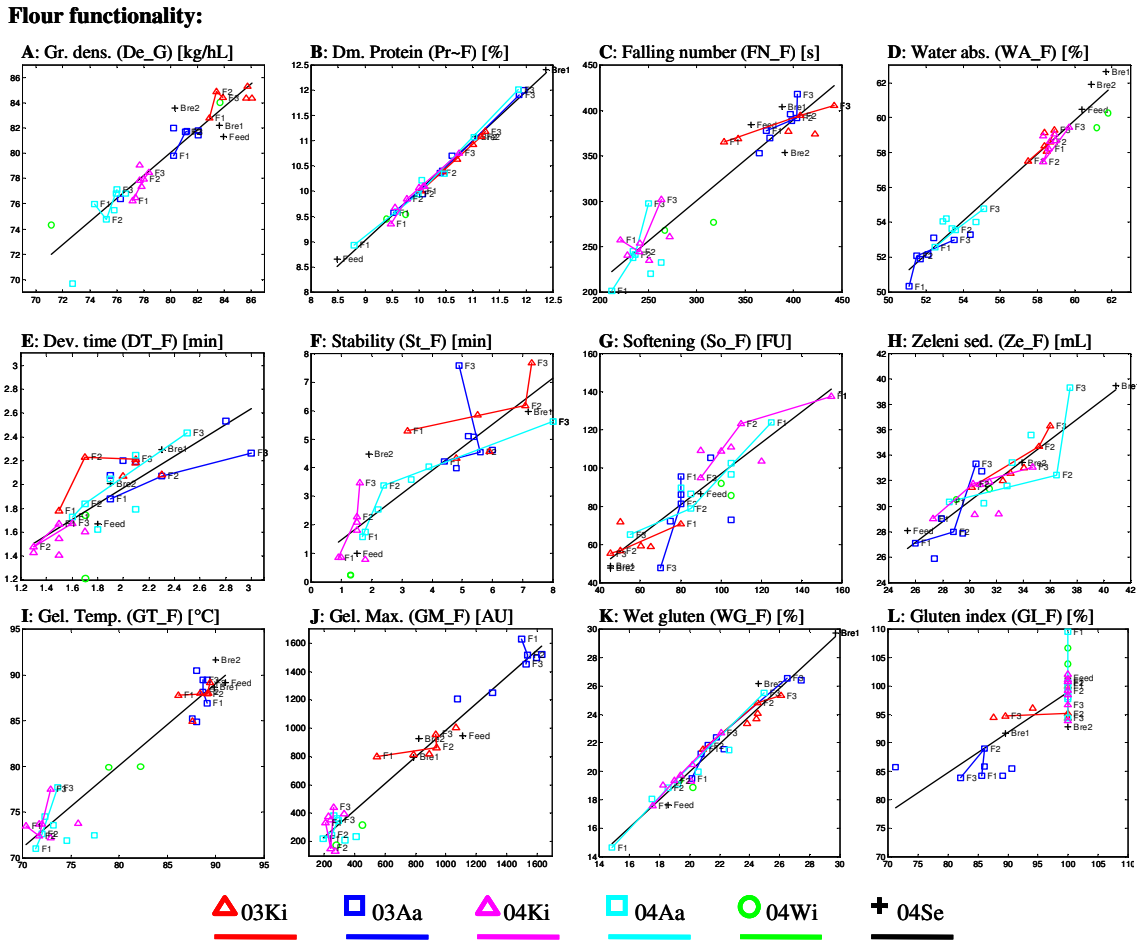


Fig. 5. Predicted vs. measured of all flour quality parameters using the MBPLS models highlighted in Table III. Samples are coloured according to origin and fractions from quality sorting are connected with lines. Legend in figure.

While flour functionality was determined by biochemical as well as rheological, i.e. physical parameters, the bread quality was determined primarily by physical parameters reflected in bread dimensions, texture profile analysis and sensory texture analysis. The relation of bread physical quality back to physicochemical measurements, i.e. spectroscopy as well as functionality, is thus quite complex and make predictions much more difficult which was demonstrated in Fig 4B with ample clarity. For a closer look at what is gained by combining the blocks

using MBPLS for prediction of bread quality, Table IV contains the similar information as Table III only regarding prediction of each individual property of bread quality. The best models with respect to both low complexity and low *RMSECV* are highlighted in the table, showing that bread mass (Ma_B), relative water loss (WL_B), TPA adhesiveness (AdFB) and sensory attributes, elasticity (ElTB), fracturability (FrTB) and dryness (DrTB) were not predicted better by adding more blocks to $\mathbf{X}_7$. By adding NIT of grain ($\mathbf{X}_1$) with the same weight as the flour quality parameters to the predictors, the bread volume, TPA hardness 1 and 2 (H1FB, H2FB) as well as the related parameters gumminess (GuFB) and chewiness (ChFB) were improved as was the hardness and adhesiveness evaluated by the sensory panel (HaTB and AhTB).

TABLE IV

Optimal number of PCs for individual bread quality parameters using the four PLS and MBPLS models with the best overall performances (abbreviations in Table II).

	$\mathbf{X}_7$: Flour functionality				$\mathbf{X}_1\mathbf{X}_7$: NIT grain + Flour func.				$\mathbf{X}_1\mathbf{X}_2\mathbf{X}_7$: NIT grain, flour + Fl.func.				$\mathbf{X}_1\mathbf{X}_2\mathbf{X}_5\mathbf{X}_7$: NIT gr., fl., NMR, Fl.func.			
	#Comp	*r*	*RMSECV*	%Y	#Comp	*r*	*RMSECV*	%Y	#Comp	*r*	*RMSECV*	%Y	#Comp	*r*	*RMSECV*	%Y
Ma_B	**4**	**0.97**	**5.0**	**94.2**	5	0.97	5.2	93.6	14	0.98	3.7	96.8	7	0.96	5.7	92.2
WL_B	**3**	**0.84**	**0.0119**	**70.3**	4	0.84	0.0120	69.8	5	0.82	0.0126	66.9	4	0.83	0.0122	69.1
Vo_B	5	0.91	176	82.2	**6**	**0.92**	**164**	**84.5**	5	0.91	169	83.7	6	0.90	184	80.6
De_B	5	0.93	0.015	86.4	**6**	**0.95**	**0.014**	**89.5**	5	0.94	0.014	88.5	8	0.94	0.015	87.6
H1FB	4	0.90	1.3	81.8	5	0.92	1.22	83.9	5	0.92	1.20	84.6	**6**	**0.92**	**1.19**	**84.8**
H2FB	4	0.92	0.76	84.7	**5**	**0.93**	**0.69**	**87.4**	6	0.94	0.69	87.4	7	0.93	0.70	87.1
AFFB	3	0.72	0.096	51.5	3	0.72	0.095	52.1	**6**	**0.74**	**0.092**	**55.1**	6	0.72	0.096	51.6
CoFB	2	0.51	0.036	26.1	5	0.55	0.036	28.0	7	0.61	0.034	33.5	**8**	**0.73**	**0.029**	**52.2**
SpFB	4	0.44	0.025	18.1	5	0.48	0.025	21.3	5	0.49	0.025	22.1	**8**	**0.66**	**0.021**	**42.5**
GuFB	4	0.93	0.64	86.7	5	0.94	0.60	88.5	**6**	**0.94**	**0.59**	**88.7**	7	0.94	0.61	88.2
ChFB	4	0.93	0.55	87.2	**5**	**0.94**	**0.51**	**88.6**	6	0.94	0.52	88.5	7	0.94	0.54	87.6
ReFB	2	0.45	0.029	20.2	5	0.46	0.030	18.8	7	0.49	0.030	17.6	**8**	**0.60**	**0.027**	**32.5**
AdFB	**2**	**0.53**	**0.80**	**28.2**	2	0.53	0.81	27.6	6	0.54	0.82	25.9	1	0.50	0.83	24.3
YeTB	4	0.80	1.9	63.9	5	0.81	1.87	64.8	**9**	**0.85**	**1.69**	**71.2**	7	0.81	1.86	65.1
ElTB	**8**	**0.84**	**1.13**	**70.8**	9	0.81	1.24	64.6	10	0.83	1.20	67.3	13	0.86	1.08	73.5
FrTB	**6**	**0.82**	**1.12**	**67.2**	6	0.81	1.16	64.8	6	0.80	1.18	63.9	5	0.78	1.23	60.6
HaTB	5	0.82	1.02	67.4	6	0.85	0.93	72.4	**7**	**0.87**	**0.88**	**75.6**	5	0.83	1.00	68.6
DrTB	**6**	**0.87**	**0.67**	**75.2**	6	0.84	0.74	69.7	13	0.86	0.72	71.9	8	0.82	0.78	66.7
CwTB	3	0.81	0.72	65.5	4	0.80	0.73	64.5	**7**	**0.84**	**0.67**	**69.9**	5	0.78	0.78	60.1
AhTB	5	0.63	0.73	35.7	**9**	**0.74**	**0.64**	**51.0**	8	0.71	0.66	48.1	1	0.69	0.66	47.0
AfTB	3	0.85	0.65	71.5	4	0.86	0.63	73.6	**5**	**0.86**	**0.62**	**74.2**	3	0.84	0.66	71.1

Further texture and sensory attributes were improved by adding NIT of flour ($\mathbf{X}_2$) to the predictors; TPA adhesive force (AFFB) and gumminess (GuFB), sensory yellowness (YlTB), hardness (HaTB), chewiness (CwTB) and aftertaste (AfTB). Explained variance of TPA cohesiveness (CoFB), springiness (SpFB), resilience (ReFB) and Hardness1 (H1FB) were improved by adding the NMR-

profiles ($\mathbf{X}_5$), thus confirming that CPMG relaxation curves of dough and bread contain information regarding texture (Engelsen et al 2001). However the instrumental spongyness represented by the cohesiveness, springiness and resilience (CoFB, SpFB, ReFB) as well as the instrumental and sensory adhesiveness parameters (AFFB, AdFB and AhTB) were not easily predicted. Only 28% to 55% of the variation in these parameters was explained in the models although the diversity is clearly caused by major differences between location/cultivar with the Aaslev/Vinjett lots being the most spongy and least adhesive (Fig. 6 G, H, I, L, M and T). In Tønning et al (submitted) location/cultivar factor was indeed significant for these attributes. The remaining attributes are predicted reasonably well considering the nature of the data. 94.2%, 70.3%, 84.5% and 89.5% of mass, relative water loss, volume and density respective could be explained (Fig. 6A-D), 84.8 to 88.6% of TPA hardness, gumminess and chewiness could be explained (Fig. 6E, F, J and K), while 67.2% to 75.5% of the sensory attributes was explained (Fig. 6N-U except T). As with the prediction of $\mathbf{Y}_1$ the rough overall modelling of all parameters in $\mathbf{Y}_2$ in Fig. 4B was inadequate to evaluate which combination of predictor blocks were actually optimal for the prediction of individual attributes. However, the flour functionality block ($\mathbf{X}_7$) was able to model bread mass, relative water loss, sensory elasticity, fracturability and dryness. Adding the NIT blocks of grain and flour ($\mathbf{X}_1$ and $\mathbf{X}_2$) an improvement of the prediction of bread volume and density as well as most of the related texture parameters was observed.

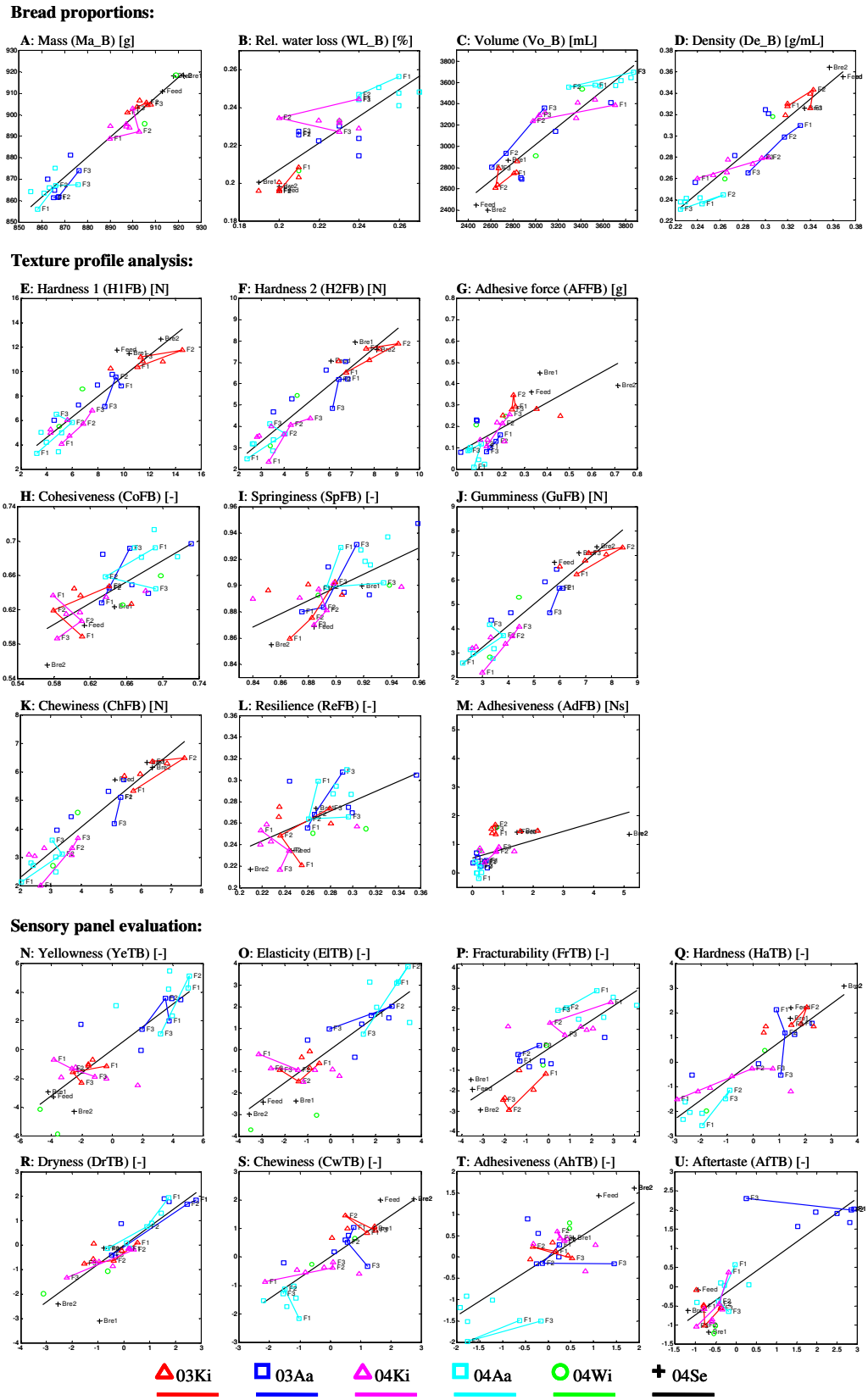


Fig. 6. Predicted vs. measured of all bread quality parameters using the MBPLS models highlighted in Table IV. Samples are coloured according to origin and fractions from quality sorting are connected with lines. Legend in figure.

In order to visualise how the bread quality data ($\mathbf{Y}_2$) are related to the flour functionality ($\mathbf{X}_7$) in these data, the regression coefficients of the PLS models corresponding to the first column in Table IV is shown in Fig, 7A and B. In Fig. 7A the regression coefficients for bread volume and all similar looking regression coefficients are shown. The volume is primarily determined by falling number, Farinograph development time and softening, Amylograph gelatinisation maximum and temperature and to some minor extent bulk grain density, protein content and gluten index. Thus the flour functionality parameters in $\mathbf{Y}_1$ ($\mathbf{X}_7$) which were predicted less well; development time, softening and gluten index, by fast spectroscopic techniques were very important in the determination of end quality in terms of volume. This explains the difficulties in determining end bread quality directly from spectroscopic analysis (Fig. 4B). The attributes with essentially similar regression coefficients are conceptually in agreement; TPA hardness, chewiness and gumminess and sensory perceived hardness, chewiness and fracturability, i.e. voluminous bread is easy to fracture and is generally softer and needs less mastication compared to more compact bread. The origin of variations in these attributes may be ascribed partially to growth year (α-amylase activity) and diversification by sorting (Farinograph development time and softening), which can be confirmed by observing the corresponding panels in Fig. 5C, E, G, I and J and Fig. 6C, E, F, J, K, P, Q and S.

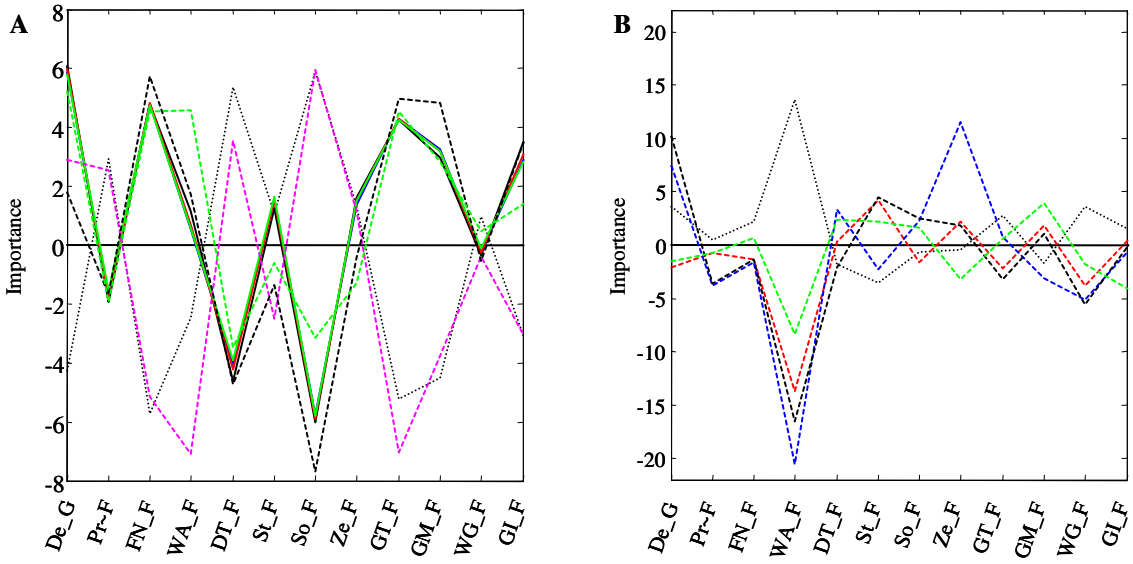


Fig. 7. Selected regression coefficients for PLS models predicting bread quality ($\mathbf{Y}_2$) from flour functionality ($\mathbf{X}_7$). **A**: **Bread volume**: ····, TPA hardness 1: —, TPA hardness 2: —, TPA gumminess: —, TPA chewiness: —, Sensory hardness: ---, Sensory chewiness: ---, Sensory fracturability: ---. **B**: **Bread mass**: ····, Sensory yellowness: ---, Sensory elasticity: ---, Sensory dryness: ---, Sensory aftertaste: ---.

In Fig. 7B the regression coefficients for bread mass and all similar looking regression coefficients are shown. The mass is (naturally) primarily determined by Farinograph water absorption. Sensory perceived dryness and elasticity as well as yellowness and aftertaste were essentially determined by the same parameter. Hence, bread with high water content appears less dry and less elastic to the sensory panel. The origin of these variations may be ascribed primarily to location/cultivar, which can be confirmed by observing the corresponding panels in Fig. 5D and Fig. 6A, N, O, R and U.

These data coherently states that being able to predict bread mass and bread volume indirectly provides information regarding key texture attributes of the bread crumb. However it is not completely clear why protein content, gluten content and Zeleny sedimentation volume is of such comparable little influence

on bread volume (Fig. 7A) as would normally be expected (Bushuk et al 1969, Veraverbeke and Delcour 2002, Wesley et al 2001). Apparently the large difference in α-amylase activity between the growth years and an unfortunate preference for weak doughs in the baking machines (Grausgruber et al 2001) may influencing the outcome of this investigation and possibly mask the expected effects of protein quantity and quality on the resulting bread products. This challenge has been discussed further in Tønning et al (submitted) while focusing on the effects of the TriQ SKNIR fractionation effects.

CONCLUSIONS

In this study PLS and MBPLS modelling was used to screen several fast multivariate tools and their combinations for prediction of flour as well as bread quality. It was shown that near-infrared transmission spectroscopy of grain and flour in combination appeared superior for predicting twelve flour functionality parameters simultaneously although NIR on flour was the single block explaining the largest number of individual flour quality parameters. Good predictions were made for protein, water absorption, wet gluten, falling number, gelatinisation temperature and gelatinisation maximum. Development time, softening and Zeleny were predicted reasonably well, while stability and gluten index were not explained well by any method. The prediction of bread quality was primarily based on flour functionality although adding NIT of grain, NIT of flour and NMR-baking profiles did improve the predictive ability of the data collected. Bread dimensions were well predicted as were instrumental texture attributes, hardness, gumminess and chewiness. Of sensory perceived attributes; yellowness, elasticity, fracturability, hardness, dryness, chewiness and aftertaste were all well determined considering the origin of the data. The MBPLS method proved useful for screening the various blocks of data and their combinations;

however screening results should be followed by more thorough investigations of performance on individual variables as shown here. The flour functionality parameters proved essential for prediction of end product quality. Especially bulk grain density, water absorption, development time, softening, falling number, gelatinisation temperature and gelatinisation maximum and gluten index were important for prediction of end product quality. By assessing the origin of the variation more inight is gained regarding possible confounding effects.

ACKNOWLEDGEMENTS

We are grateful to the careful laboratory work by Jette Ryaa Nielsen and the in-house sensory panel at University of Aarhus, Faculty of Agricultural Sciences, Department of Food Science and Lisbeth T. Hansen at University of Copenhagen, Faculty of Life Sciences, Department of Food Science. Bo Löfqvist from BoMill AB is gratefully acknowledged for conducting fractionation of our starting material and for valuable discussions. We thank Cerealia Mills, Vejle, Denmark for facilitating laboratory help, space and equipment. We are grateful to Lars Munck and Åse Hansen University of Copenhagen, Faculty of Life Sciences, Department of Food Science and Johannes Ravn Jørgensen and Bernd Wollenweber at University of Aarhus, Faculty of Agricultural Sciences, Department of Genetics and Biotechnology for help and enlightening discussions. For financial support we are greatly thankful to the Interreg IIIA programme between Fyns Amt, Odense, Denmark and Technologie-Region K.E.R.N. Rendsburg, Germany and to University of Copenhagen, Faculty of Life Sciences for the ph.d. scholarship to the first author, and to the project FFS05-9: Build Your Food sponsored by the Ministry of Food, Agriculture and fisheries.

LITERATURE CITED

AACC Interntional. 2000. Approved Methods of the American Association of Cereal Chemists, 10th ed. Methods 22-10, 38-12, 56-61A. The Association: St. Paul, Minnesota, USA.

Brás, L. P., Bernardino, S. A., Lopes, J. A., and Menezes, J. C. 2005. Multiblock PLS as an approach to compare and combine NIR and MIR spectra in calibrations of soybean flour. Chemometr. Intell. Lab. 75:91-99.

Berglund, A., and Wold, S. 1999. A serial extension of multiblock PLS. J. Chemometr. 13:461-471.

Bushuk, W., Briggs, K. G., Shebeski, L. H. 1969. Protein quantity and quality as factors in the evaluation of bread wheats. Can. J. Plant Sci., 49:113-122.

Callaghan, P. T. 1991. Principles of Nuclear Magnetic Resonance Microscopy. Clarendon Press: Oxford.

Carr, H. Y., and Purcell, E. M. 1954. Effect of diffusion on free precession in nuclear magnetic resonance experiments. Phys. Rev. 94:630-638.

Choi, S. W., and Lee, I. B. 2005. Multiblock PLS-based localized process diagnosis. J. Process Contr. 15:295-306.

Delwiche, S. R., and Weaver, G. 1994. Bread Quality of Wheat-Flour by Near-Infrared Spectrophotometry - Feasibility of Modelling. J. Food Sci. 59:410-415.

Delwiche, S. R., Graybosch, R.A., and Peterson, C.J. 1998. Predicting protein composition, biochemical properties, and dough-handling properties of hard red winter wheat flour by near-infrared reflectance. Cereal Chem. 75:412-416.

Dowell, F. E., Maghirang, E. B., Xie, F., Lookhart, G. L., Pierce, R. O., Seabourn, B. W., Bean, S. R., Wilson, J. D., and Chung, O. K. 2006. Predicting wheat quality characteristics and functionality using near-infrared spectroscopy. Cereal Chem. 83:529-536.

Dyrby, M., Nørgaard, L., and Engelsen, S. B. 2002. Chemometric Quantitation of the Active Substance (Containing C≡N) In a Pharmaceutical Tablet Using Near-Infrared (NIR) Transmittance and NIR FT-Raman Spectra. Appl. Spectrosc., 56:579-585.

Engelsen, S. B., Jensen, M. K, Pedersen, H. T., Nørgaard, L., and Munck, L. 2001. NMR-baking and multivariate prediction of instrumental texture parameters in bread. J. Cereal Sci. 33:59-69.

Felicio, C. C., Brás, L. P., Lopes, J. A., Cabrita, L., and Menezes, J. C. 2005. Comparison of PLS algorithms in gasoline and monitoring with MIR and NIR. Chemometr. Intell. Lab. 78:74-80.

Fontaine, J., Schirmer, B., and Horr, J. 2002. Near-Infrared Reflectance Spectroscopy (NIRS) enables the fast and accurate prediction of essential amino acid contents. 2. Results for wheat, barley, corn, triticale, wheat bran/middlings, rice bran, and sorghum. J. Agri. Food Chem. 50:3902-3911.

Frank, I. E., Feikema, J., Constantine, N., and Kowalski, B. R. 1984. Prediction of product quality from spectral data using the partial least squares method. J. Chem. Inf. Comp. Sci. 24:20-24.

Geladi, P., and Kowalski, B.R. 1986. Partial least-squares regression: A Tutorial. Anal. Chim. Acta 185:1-17

Grausgruber, H., Kreuzmayr, A. E., and Ruckenbauer, P. 2001. Evaluation of the breadmaking quality of Austrian-grown wheats using an automatic home-bakery. Cereal Res. Commun. 29:421-428.

Hansen, B., and Hansen, Å. 1992. Test baking of bread by household baking machine. Food Sci. Tech. 25:585-587.

Hansen, B., and Hansen, Å. 1993. Erratum. Test baking of bread by household baking machine. Food Sci. Tech. 26:181.

Hotelling, H. 1933. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 24:417-441, 498-520.

Hrušková, M, Bednářová, M., and Šmejda, P. 2004. Prediction of rheological parameters of dough by NIR spectral analysis of wheat flour. Chem. Listy 98:423-431.

ICC. 1998. Standard Methods of the International Association for Cereal Chemistry (ICC). Standrard No. 107/1, 115/1, 116/1, 126/1, 155. Verlag Moritz Schäfer: Detmold, Germany.

Kourti, T. 2006. Process analytical technology beyond real-time analyzers: The role of multivariate analysis. Crit. Rev. Anal. Chem. 36:257-278.

Martens, H., Anderssen, E., Flatberg, A., Gidskehaug, L. H., Hoy, M., Westad, F., Thybo, A., and Martens, M. 2005. Regression of a data matrix on descriptors of both its rows and of its columns via latent variables: L-PLSR. Comput. Stat. Data An. 48:103-123.

Martens, H., Nielsen, J. P., and Engelsen, S. B. 2003. Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. Anal. Chem. 75:394-404.

Martens, H, and Stark, E. 1991. Extended multiplicative signal correction and spectral interference subtraction – new processing methods for near infrared spectroscopy. J. Pharmaceut. Biomed. 9:625-635.

Martens, M., Bredie, W. L. P., Martens, H. 2000. Sensory profiling data studied by partial least squares regression. Food Qual. Prefer. 11:147-149.

Meiboom, S., and Gill, D. 1958. Modified spin-echo method for measuring nuclear relaxation times. Rev. Sci. Instrum. 29:688-691.

Micklander, E., Peshlov, B., Purslow, P. P., and Engelsen, S. B. 2002. NMR cooking: Monitoring the changes in meat during cooking by low-field 1H-NMR, Trends Food Sci. Tech. 13:341-346.

Miralbés, C. 2003. Prediction chemical composition and Alveograph parameters on wheat by near-infrared transmittance spectroscopy. J. Agri. Food Chem. 51:6335-6339.

Miralbés, C. 2004. Quality control in the milling industry using near infrared transmittance spectroscopy. Food Chem. 88:621-628.

Mortensen, M., Thybo, A. K., Bertram, H. C., Andersen, H. J., and Engelsen, S. B. 2005. Cooking effects on water distribution and mobility in potatoes using NMR relaxation. J. Agri. Food Chem. 53:5976-5981

Munck, L. 2005. The revolutionary aspect of exploratory chemometric technology, Narayana Press, Gylling: Denmark.

Osborne, B.G., and Douglas, S. 1981. Measurement of the Degree of Starch Damage in Flour by Near-Infrared Reflectance Analysis. J. Sci. Food Agri. 32:328-332.

Peltonen, J., and Salovaara, H. 1991. Experiences of An Automatic Small-Scale Home Bakery in Test Bakings of 6 Spring Wheat-Varieties. J. Agr. Sci. Finland 63:131-135.

Qin, J.S, Valle, S., and Piovoso, M. J. 2001. On unifying multiblock analysis with application to decentralized process monitoring. J. Chemometr. 15:715-742.

Tønning, E., Thybo, A. K., Pedersen, L., Munck, L., Hansen, Å., Engelsen, S. B., and Nørgarrd, L. Bulk quality diversification of organic wheat by single-kernel near-infrared (SKNIR) sorting. J. Cereal Sci., submitted.

USFDA. 2004. Guidance for industry. PAT – A framework for innovative pharmaceutical development, manufacturing and quality assurance. US Department of Health and Human Services, Center for Drug Evaluation and Research (CDER), Rockville, MD 20857, United States. Published online at http://www.fda.gov/cder/guidance/6419fnl.pdf. June 2007.

van den Berg, F., Povlsen, V., Thybo, A. and Bro, R. 2001. Multi-block methods for exploratory data mining in food technology. Department of Food Science, University of Copenhagen. Published online at http://www.models.life.ku.dk/courses/MBtoolbox/MBIntroEn010331.htm, June 2007.

Veraverbeke, W. S., and Delcour, J. A. 2002. Wheat protein composition and properties of wheat glutenin in relation to breadmaking functionality. Crit. Rev. Food Sci. 42:179-208.

Vivien, M., and Sabatier, R. 2003. Generalized orthogonal multiple co-inertia analysis(-PLS): new multiblock component and regression methods. J. Chemometr. 17:287-301.

Vivien, M., Verron, T., and Sabatier, R. 2005. Comparing and predicting sensory profiles from NIRS data: use of the GOMCIA and GOMCIA-PLS multiblock methods. J. Chemometr. 19:162-170.

Wesley, I. J., Larroque, O., Osborne, B. G., Azudin, N., Allen, H., and Skerritt, J. H. 2001. Measurement of gliadin and glutenin content of flour by NIR spectroscopy. J. Cereal Sci. 34:125-133.

Westerhuis, J. A., and Coenegracht, P. M. J. 1997. Multivariate modelling of the pharmaceutical two-step process of wet granulation and tableting with multiblock partial least squares. J. Chemometr. 11:379-392.

Westerhuis, J. A, Kourti, T., and Macgregor, J. F. 1998. Analysis of multiblock and hierarchical PCA and PLS models. J. Chemometr. 12:301-321.

Westerhuis, J. A, and Smilde, A. K. 2001. Deflation in multiblock PLS. J. Chemometr. 15:485-493.

Wold, S., Martens, H., and Wold, H. 1983. The multivariate calibration-problem in chemistry solved by the PLS method. Lect. Notes Math. 973:286-293.

Wold, S., Esbensen, K., and Geladi, P. 1987. Principal Component Analysis. Chemometr. Intell. Lab. 2:37-52.

Zachariassen, C. B., Larsen, J., van den Berg, F., and Engelsen, S. B. 2005. Use of NIR spectroscopy and chemometrics for on-line process monitoring of ammonia in Low Methoxylated Amidated pectin production. Chemometr. Intell. Lab. 76:149-161.

Zwingelberg, H., and Brümmer, J.-M. 1990. Backautomaten in Mühlenlaboratorien. Getreide Mehl Brot 44:142-147.

Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 188 (2007) 10–23
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

A novel improved method for analysis of 2D diffusion–relaxation data—2D PARAFAC-Laplace decomposition

Erik Tønning [a,*], Daniel Polders [b], Paul T. Callaghan [c], Søren B. Engelsen [a]

[a] *Quality & Technology, Department of Food Science, Faculty of Life Sciences, University of Copenhagen, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark*[1]

[b] *Laboratory for Biophysics, Wageningen University and Research Centre, Dreijenlaan 1, 6703 HA Wageningen, The Netherlands*[2]

[c] *MacDiarmid Institute for Advanced Materials and Nanotechnology, Victoria University of Wellington, New Zealand*[3]

Received 26 March 2007

Abstract

This paper demonstrates how the multi-linear PARAFAC model can with advantage be used to decompose 2D diffusion–relaxation correlation NMR spectra prior to 2D-Laplace inversion to the T_2–D domain. The decomposition is advantageous for better *interpretation* of the complex correlation maps as well as for the *quantification* of extracted T_2–D components. To demonstrate the new method seventeen mixtures of wheat flour, starch, gluten, oil and water were prepared and measured with a 300 MHz nuclear magnetic resonance (NMR) spectrometer using a pulsed gradient stimulated echo (PGSTE) pulse sequence followed by a Carr–Purcell–Meiboom–Gill (CPMG) pulse echo train. By varying the gradient strength, 2D diffusion–relaxation data were recorded for each sample. From these double exponentially decaying relaxation data the PARAFAC algorithm extracted two unique diffusion–relaxation components, explaining 99.8% of the variation in the data set. These two components were subsequently transformed to the T_2–D domain using 2D-inverse Laplace transformation and quantitatively assigned to the oil and water components of the samples. The oil component was one distinct distribution with peak intensity at $D = 3 \times 10^{-12}$ m^2 s^{-1} and $T_2 = 180$ ms. The water component consisted of two broad populations of water molecules with diffusion coefficients and relaxation times centered around correlation pairs: $D = 10^{-9}$ m^2 s^{-1}, $T_2 = 10$ ms and $D = 3 \times 10^{-13}$ m^2 s^{-1}, $T_2 = 13$ ms. Small spurious peaks observed in the inverse Laplace transformation of original complex data were effectively filtered by the PARAFAC decomposition and thus considered artefacts from the complex Laplace transformation. The oil-to-water ratio determined by PARAFAC followed by 2D-Laplace inversion was perfectly correlated with known oil-to-water ratio of the samples. The new method of using PARAFAC prior to the 2D-Laplace inversion proved to have superior potential in analysis of diffusion–relaxation spectra, as it improves not only the interpretation, but also the quantification.

Keywords: DRCOSY; PARAFAC; Laplace inversion; Diffusion; Relaxation; Correlation spectroscopy; NMR; PGSTE; Dough; Water; Oil

1. Introduction

Characterisation of water and fat components in food is of prime importance due to its modulation of important properties such as taste, texture, oxidation and shelf life. Diffusion correlated NMR relaxometry is a unique technique for characterisation of dynamics, compartmentalisation and phases of fat and water in food, as it is able to measure complex solid or semi-solid food matrices such as meat, cheese, dough and bread. This correlation technique, in which 2D-Laplace inversion NMR is used to provide a map in T_2–D space, is used to analyse the complex multi-exponential behaviour of the relaxation and diffusion rates in heterogeneous systems. It enables us to obtain a plot that is easy to interpret and separates components of a system via their dynamics, revealing additional informa-

[1] http://www.models.life.ku.dk.
[2] http://www.bip.wur.nl/UK.
[3] http://www.macdiarmid.ac.nz/nmr.

* Corresponding author. Fax: +45 35283245.
E-mail address: ert@life.ku.dk (E. Tønning).

1090-7807/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2007.05.018

tion by correlating these molecular motions when compared with a 1D technique [1–3].

2D diffusion–relaxation data are double-exponentially decaying landscapes and thus second order data structures. Data from series of complex samples can with advantage be analysed using multi-way chemometric methods, as these will provide for unique resolution of pure component landscapes. Multi-way analysis has successfully been applied in several chemical fields including fluorescence emission–excitation spectroscopy [4] and 2D NMR spectroscopy [5,6]. The key issue in multi-way analysis is to have access to boxes of data rather than tables of data. Usually, a single spectrum is recorded for each sample. Data for several samples are then gathered in a matrix/table. If, instead, the data are recorded as a function of two variables (e.g. residual magnetisation as a function of time, yielding relaxation, and as a function of magnetic field gradient, yielding diffusion), then the data from one sample are contained in a matrix. For several samples a box of data is obtained. Such multi-way data can be modelled with specialized tools that take particular advantage of the data structure. Most notably, the so-called PARAFAC model [7] is an interesting alternative to traditional data analysis tools, because it allows resolving complex mixture measurements into the pure single-component spectra. The advantage of PARAFAC in this context is its ability to provide unique solutions to data that are approximately multi-linear. T_2–D weighted relaxation data is one example of trilinear data. These can be decomposed by PARAFAC into a few pure and unique physico-chemical components with exactly the same data structure as the original data. A subsequent T_2–D Laplace inversion of the resolved components will then provide the T_2–D distribution profile of the pure components. The PARAFAC algorithm thus works as a filter, extracting only the systematic variation from a coherent set of DRCOSY recordings, while leaving out the non-systematic variation in the residuals. The combined method constitutes a significant improvement to the complex T_2–D Laplace inversion of individual samples in which the researcher has no objective means of assessing if individual peaks represent physico-chemical components or artefacts that appear due to the ill-conditioned problem and unfortunate choice of regularisation. A conceptual sketch of the new composite method demonstrated in this paper is shown in Fig. 1.

2. Theory

The correlated measurement of diffusion and spin relations by NMR requires the use of an r.f. and magnetic field gradient pulse train which "encodes" for both parameters on the same nuclear spin magnetisation. The separate encoding methods for diffusion and relaxation are described in detail elsewhere [8]. The combined DRCOSY method uses a pulsed gradient stimulated echo (PGSTE) [9,10] followed by a Carr–Purcell–Meiboom–Gill (CPMG) pulse echo train [11,12]. During that CPMG train the spin magnetisation signal, M, relative to the initial echo amplitude, M_0, is sampled, both as a function of time, t, during the train, and gradient strength, q^2 applied in the PGSTE sequence before the train. This the signal is acquired in a 2D (t, q^2) space as:

$$\frac{M(t,q^2)}{M_0} = \sum p(D,T_2)\mathrm{e}^{-q^2D(\Delta-\frac{\delta}{3})}\mathrm{e}^{\frac{-t}{T_2}} + \mathrm{e}(t,q^2) \quad \text{and} \quad q = G\gamma\delta, \tag{1}$$

where Δ is the diffusion observation time, γ is the gyromagnetic ratio, δ is the gradient duration and G is the gradient strength. Eq. (1) thus assumes a distribution of diffusion coefficients, D, and relaxation times, T_2, with joint probability, p [13].

In order to obtain the distribution p, the experimental data must be inverted using 2D-Laplace inversion. This is done according to Song et al. (2002) [14] by considering a the discrete matrix form of Eq. (1):

$$\mathbf{M} = \mathbf{K_1}\mathbf{X}\mathbf{K_2'} + \mathbf{E} \tag{2}$$

where $\mathbf{K}_1$ and $\mathbf{K}_2$ are the known matrices of the exponentials in Eq. (1) for the observation time and gradient strengths used while *choosing* a discrete number of relaxation times and diffusion coefficients in a specified range. Thus, the window of observation and resolution is chosen by the investigator and the applicability (robustness) limits of the algorithm. $\mathbf{X}$ is the unknown T_2–D distribution matrix extracted by minimising:

$$\chi^2 = \|\mathbf{M} - \mathbf{K_1}\mathbf{X}\mathbf{K_2'}\|^2 + \alpha\|\mathbf{X}\|^2 \tag{3}$$

where α is a regularisation factor set by the user depending on the desired smoothness of the result and $\|\cdot\|$ is the Frobenius norm. Choosing an appropriate α is not straightforward; however, when following the usual guidelines [14], in which α is adjusted to just minimize χ^2, the results are often readily interpretable. In contrast to the PARAFAC model, the Laplace inversion problem is ill-conditioned and depends on the algorithm used for regularisation and the set of parameters used. Thus, caution must be taken when interpreting and quantifying peaks in the T_2–D spectra.

In this work we aim to investigate if the inherent labile nature of the Laplace inversion procedure can be improved by resolving (filtering) the unique components using the PARAFAC algorithm [7] prior to the 2D-Laplace transformation of the data. Using parallel factor analysis (PARAFAC) we stack the 2D landscapes of the different samples in a 3D array, $\underline{\mathbf{M}}$, of the size: $I \times J \times K$. I is the number of samples, J is the number of gradient steps and K is the number of time points. If the data is tri-linear, each observation point, m_{ijk} in $\underline{\mathbf{M}}$, can be described uniquely as:

$$m_{ijk} = \sum_{f=1}^{F} a_{if}b_{jf}c_{kf} + e_{ijk}, \quad i = 1,\ldots,I; \quad j = 1,\ldots,J; \quad k = 1,\ldots,K \tag{4}$$

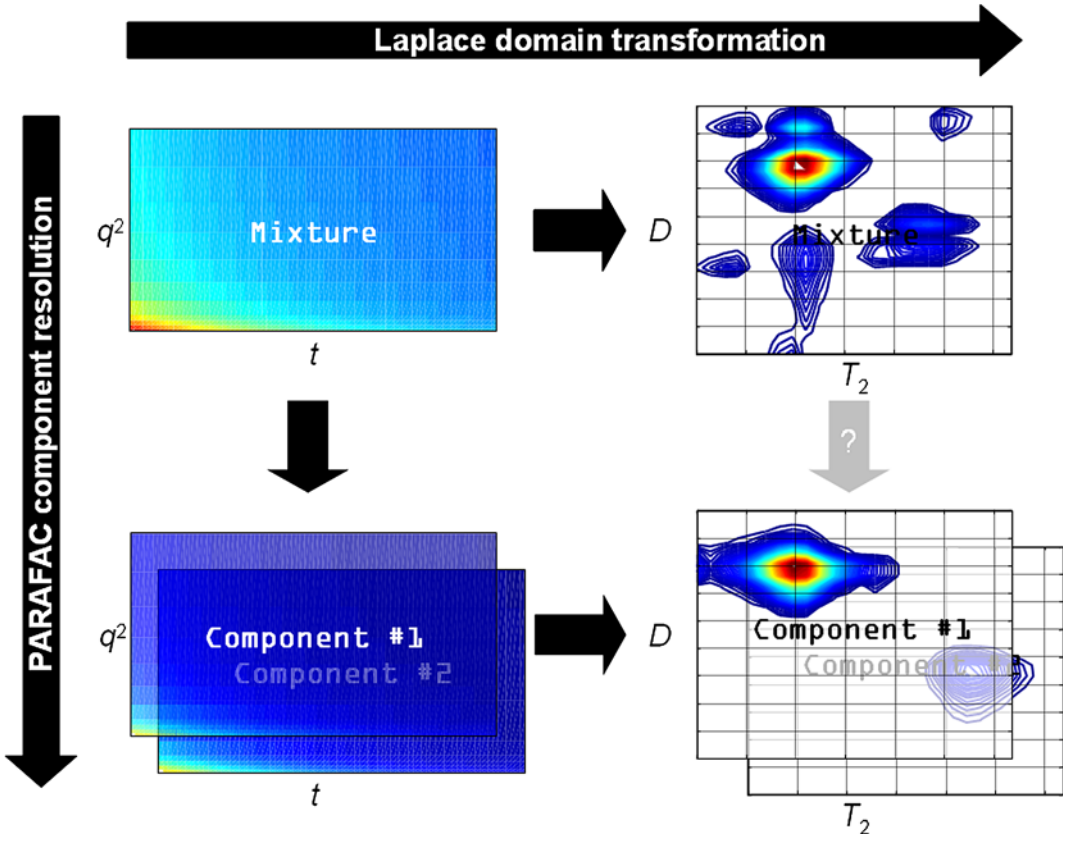


Fig. 1. Concept scheme of the method demonstrated. Raw 2D diffusion–relaxation NMR data (upper left) can be directly transformed into the T_2–D domain (upper right) by Laplace inversion or as suggested here via spectral decomposition by PARAFAC (lower left) to unique T_2–D PARAFAC-Laplace components (lower right).

where F is the number of PARAFAC components, $\mathbf{a}_f$, $\mathbf{b}_f$ and $\mathbf{c}_f$ are the PARAFAC scores of lengths I, J and K, respectively, and a_{if} is the ith element of $\mathbf{a}_f$, b_{jf} is the jth element of $\mathbf{b}_f$ and c_{kf} is the kth element of $\mathbf{c}_f$. A graphical representation of the PARAFAC model with two components is found in Fig. 2, where $\underline{\mathbf{M}}$ is the original box of data. Mode 1 is samples, mode 2 is relaxation time, t, and mode 3 is the gradient strength, q^2. PARAFAC decomposes the data into two unique components consisting of three loading vectors, one for each mode, and a residual matrix, $\underline{\mathbf{E}}$. By taking the outer product of loading 2, $\mathbf{b}_f$, and loading 3, $\mathbf{c}_f$, of each component the components are now represented by a 2D-loading spectrum with an associated sample loading, $\mathbf{a}_f$, holding the concentrations of the components of in each sample. The sample loading vectors are also termed score vectors. Thus each sample is now decomposed into a weighted sum of unique 2D diffusion–relaxation spectra and a residual matrix. The weights are the sample scores in mode 1.

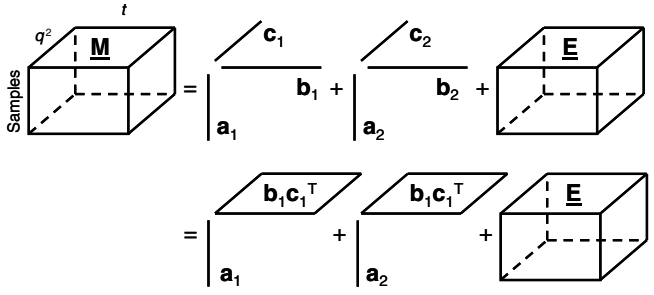


Fig. 2. Graphical representation of a 3-way array, $\underline{\mathbf{M}}$, decomposed by PARAFAC into two unique components and a residual, $\underline{\mathbf{E}}$. See text for thorough explanation.

Eq. (4) is completely unaffected by the underlying distributions in Eq. (1) of the pure physico-chemical components, as the PARAFAC algorithm is not restricted by the mathematical relationship within the components. PARAFAC simply extracts independently varying additive components. The prerequisite for this extraction is, that a number of samples is recorded in a manner by which the components of interest are purposely varied either by design or by natural variation.

The aim of this investigation is to demonstrate that PARAFAC in practice is able to resolve pure 2D diffusion–relaxation components on a set of samples containing the same components, but in different proportions. When this is achieved and the quantification is thus in place, interpretation is straightforward by subsequent application of the 2D Laplace inversion to the pure 2D diffusion–relaxation components. The hypothesis is that application of PARAFAC will facilitate a more direct interpretation and robust application of the 2D Laplace inversion. The

particular experimental system used comprises wheat doughs. This system was chosen, because it provides sufficient complexity in the T_2–D spectrum and because of the inherent interest in developing robust 2D-inverse Laplace methods for the food industry.

3. Experimental

3.1. Materials

Five wheat flour samples from a Danish/German field experiment were used. They were chosen for their diverse quality with respect to baking quality and functionality. An even greater diversification of samples in this investigation was obtained by mixing the following ingredients: wheat starch (Unmodified, Sigma CAS 9005-25-8), wheat gluten (Sigma, CAS 8002-80-0), commercially available soy oil and commercially available high-grade wheat flour available in New Zealand (trademark: Champion).

3.2. Characterisations

An extensive characterisation of the Danish/German wheat flours was performed for other purposes and will be elaborated elsewhere. The mixing property with regards to water uptake was investigated using a Farinograph according to ICC standard No. 115/1. The commercially available flour was characterised using a Foss NIRS (near-infrared reflectance spectroscopy) Bench Analyser with an in-house calibration at Weston Milling, Lower Hutt, Wellington for estimation of several parameters including moisture and Farinograph water absorption. The water content of the flours as well as the starch and the gluten were determined by gravitational method prior to mixing the doughs. Two gram of the materials were dried at 130 °C for 1.5 h and scaled before and after according to ICC standard No. 110/1 (Table 1).

3.3. Dough preparations

Twelve samples were prepared by mixing wheat starch, wheat gluten and soy oil following a full factorial design with three centre points (Fig. 3) and with the commercially available wheat flour as filler to a 43 g dm sample, i.e. equal to 50 g sample at 14% moisture. The factorial ingredients were used in three levels each: starch: 0, 10.0 and 20.0 g dm; gluten: 0, 3.5 and 7.0 g dm; oil: 0, 1.0 and 2.0 g (Fig. 3). The dry matter mass of commercial flour was thus dependent on the levels of pure ingredients and varied between 14.0 and 43.0 g dm. To all mixtures 30 g of water was added, thus assuming an average of 60% water uptake based on 14% moisture in all the mixtures. The sum of the moisture contents of the ingredients and the added water in the samples added up to 37 g in all mixtures equal to 46% of the total mass (Table 1). The mixed samples were prepared and recorded in random order. The Danish/German wheat flour samples were included for natural variation in the experimental setup (Fig. 3). They were only mixed with water according to their water absorption and actual moisture content (Table 1).

A 50MDD Laboratory Mixer, Lincoln, New Zealand was used for preparing the doughs. In the running mixer the flour samples and, respectively, the mixture samples were added to the mixing chamber. The ingredients of the mixture samples were added consecutively in the following order: flour, starch, gluten and soy oil. After tempering at 34 °C for 2 min the dough preparation was initiated by adding water (34 °C) in amounts corresponding to the water absorption capacity and the energy counter was simultaneously reset. The doughs were mixed for 49–259 s depending on the dough consistency development until the energy input reached 10.0 Wh/kg dough. The aim of the procedure was to add an equal amount of work into each sample and thus to produce samples in a uniform and reproducible way.

3.4. NMR recordings

Immediately after preparation a small amount of dough was inserted bit by bit into a 5 mm wide NMR tube using a piston rod to pack the material while avoiding air bubbles to form. The tube was filled up to 3 cm in order to fully cover the sampling area of the tube. The tube was inserted into a Bruker Avance 300 System fitted with a Bruker 36 T m^{-1} gradient coil, capable of applying a strong specific perturbation of the magnetic field along the z direction. The machine was operated from a UNIX PC running xwinnmr version 3.6.

A DRCOSY pulse program was written consisting of a pulsed gradient stimulated echo (PGSTE) followed by a Carr–Purcell–Meiboom–Gill (CPMG) echo train (Fig. 4). The $\pi/2$ and π hard pulses were applied for 5.2 and 10.4 μs, respectively. The PGSTE was initiated by a $\pi/2$ hard pulse, initiating a free induction decay (FID) followed by a magnetic field gradient pulse, two consecutive $\pi/2$ hard pulses and a magnetic field gradient pulse of exactly the same size as the first after which a stimulated echo occurs depending on the size of the gradient. The peak intensities of echoes were then recorded during a CPMG sequence.

In 25 consecutive PGSTE+CPMG runs the gradient strength was varied from 0 to 12.96 T m^{-1} in 25 approximately exponentially spaced steps with constant diffusion observation time, $\Delta = 20.00$ ms and gradient duration, $\delta = 2.00$ ms with ramp times of 500 μs in 10 steps, i.e. ramp up time of 500 μs, stable time of 1500 μs and ramp down time of 500 μs. The q-encoding gradient pulses were flanked by delays of 500 μs immediately before and after the gradient. The latter was applied specifically to allow for the ring down of induced eddy currents in the surrounding metals before applying the 2nd $\pi/2$ hard pulse, the spin-conserving pulse. During the z-storage time of 16,489.6 μs before the 3rd $\pi/2$ hard pulse a crusher gradient (homospoil) was ramped in 4 steps of 50 μs to 5% of max-

Table 1
Twelve mixtures named by their contents of starch (S), gluten (G), fat (F) denoted by 0, 1 and 2 for levels zero, medium and high, respectively

Sample #	ID	Quantities of ingredients (% moisture)							Mixing time	Totals		
		Starch (11.2%) (g)	Gluten (7.9%) (g)	Fat (—) (g)	Flour (12.9%) (g)	Moist. in mixture (%)	Water upt. (14% moist.) (%)	Water added (g)		Sample mass (g)	Water (%)	Added fat (%)
S01	S0G0F0	0.0	0.00	0.0	49.4	12.9	60.0	30.6	01:05	80.0	46.3	0.00
S02	S0G2F0	0.0	7.60	0.0	41.3	12.1	60.0	31.1	00:49	80.0	46.3	0.00
S03	S2G0F0	22.5	0.00	0.0	26.4	12.1	60.0	31.1	01:55	80.0	46.3	0.00
S04	S2G2F0	22.5	7.60	0.0	18.4	11.3	60.0	31.5	01:12	80.0	46.3	0.00
S05	S0G0F2	0.0	0.00	2.0	47.1	12.3	60.0	30.9	01:57	80.0	46.3	2.50
S06	S0G2F2	0.0	7.60	2.0	39.0	11.6	60.0	31.4	01:14	80.0	46.3	2.50
S07	S2G0F2	22.5	0.00	2.0	24.1	11.6	60.0	31.4	04:19	80.0	46.3	2.50
S08	S2G2F2	22.5	7.60	2.0	16.1	10.8	60.0	31.8	01:42	80.0	46.3	2.50
S09	S2G2F2	22.5	7.60	2.0	16.1	10.8	60.0	31.8	02:19	80.0	46.3	2.50
S10	S1G1F1	11.3	3.80	1.0	32.7	11.8	60.0	31.2	01:27	80.0	46.3	1.25
S11	S1G1F1	11.3	3.80	1.0	32.7	11.8	60.0	31.2	01:27	80.0	46.3	1.25
S12	S1G1F1	11.3	3.80	1.0	32.7	11.8	60.0	31.2	01:34	80.0	46.3	1.25
S13	03KCF3	—	—	—	50.1	14.1	58.9	29.4	01:24	79.5	45.9	0.00
S14	03AVF3	—	—	—	50.3	14.5	53.5	26.5	00:55	76.8	44.0	0.00
S15	04KCF1	—	—	—	49.8	13.6	58.6	29.5	01:21	79.3	45.8	0.00
S16	04AVF1	—	—	—	49.8	13.7	52.5	26.4	01:17	76.3	43.6	0.00
S17	04APRf	—	—	—	49.8	13.6	61.2	30.8	01:21	80.6	46.7	0.00

Flour is added to obtain a 43 g dm sample and water is added assuming a 60% water uptake (based on 14% water) for all mixtures. Five additional samples flour of different origin, water content and water uptake is also included. Names refer to year (03/04), place (K, Kiel; A, Aarslev), cultivar (C, Carbo; V, Vinjet; P, Pentium) and treatment (F1, fraction 1; F3, fraction 3; Rf, Reference).

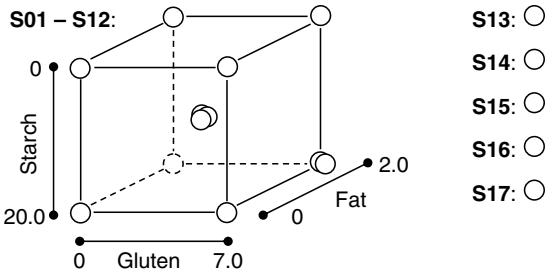


Fig. 3. The experimental design. Twelve mixture samples (S01–S12) in a full factorial design of experiment with three factors; Starch at 0, 10.0 and 20.0 g; gluten at 0, 3.5 and 7.0 g; soy oil at 0, 1.0 and 2.0 g. There are three centre points (S1G1F1) and the sample with high levels in all factors (S2G2F2) is represented twice. Five additional samples (S13–S17) are indicated. See Table 1 for details.

imum gradient strength (i.e. 1.8 T m^{-1}) and kept for 1.0 ms before being ramped down again. By destroying unwanted transverse magnetisation the crusher effectively removed potential interference from the spin echo of the initial r.f. pulse-FID generated by the second hard pulse and thus also the FID from the second r.f. pulse that would otherwise become an echo after the third r.f. pulse. In order for the gradient to stabilise, i.e. reach steady state, ten dummy gradient pulses as described above, but without the hard pulses, were applied before the initial π/2 hard pulse.

The spin echo signal from the PGSTE pulse sequence was recorded during the subsequent CPMG pulse sequence with a time delay of 100 μs between the π hard pulses (10.4 μs) and the echo centres, thus the echo time, $\tau = 105.2$ μs. Three points in every second echo centre were recorded and subsequently averaged. Fourthousand and ninety-six even echoes were acquired during the 1.724 s CPMG pulse sequence. Four scans were made for every gradient step with a repetition delay of 1.977 s. The entire DRCOSY pulse sequence would thus run for 398.9 s.

3.5. Data processing

Data processing was performed using commercially available software packages, Prospa V2.0.12, Magritek, Wellington, New Zealand and Matlab Version 6.5 release 13, The MathWorks, Inc. Conversion and pre-processing of the data was made in Prospa as were the 2D-inverse Laplace transformations, while PARAFAC was run in Matlab using the *N*-way toolbox [15] from www.models.life.ku.dk. For every sample the resulting matrix of $25 \times (3 \times 4096)$ real and imaginary recordings was imported into Prospa. The three points in every echo were averaged and the real and imaginary recordings were summed. Since the DRCOSY pulse sequence does not produce meaningful data at zero gradient strength, the first data row was removed. The first column representing the initial echo recording was also removed due to deviating non-exponential values in the first of the three points recorded. Following this procedure the total data set was reduced to the dimensions: 17 samples × 24 gradient steps × 4095 acquisition times. As there was no internal standard and no control of the actual sample mass in the measured volume of the NMR tube, all samples were normalised with maximum intensity, i.e. the intensity of the first acquisition point assuming similar amounts of fast relaxation components in the samples.

For every gradient strength, G_i, the gradient axis values were calculated by:

$$\text{gradaxis}_i = q^2\left(\Delta - \frac{\delta}{3}\right) = (G_i\gamma\delta)^2\left(\Delta - \frac{\delta}{3}\right) \tag{5}$$

with the gyromagnetic ratio $\gamma = 2.675 \times 10^9$ s^{-1} T^{-1} and $G_{\min} = 0.216$ T m^{-1} and $G_{\max} = 12.96$ T m^{-1} the gradient axis thus spans from 0.258×10^9 to 929.447×10^9 s m^{-2}. The time axis associated with T_2 relaxation consisted of equidistant time point for every 420.8 μs corresponding to every 2nd echo beginning at time point 7.8416 ms corresponding to the actual duration of the two gradient pulses and the disregarded first echo. These axes were required when performing the 2D-inverse Laplace transformations in Prospa of the raw sample recordings as well as the PARAFAC components and residuals (see below).

A two-component PARAFAC model was calculated using non-negativity constraints in all three modes. For *each component* in *each sample* the three modes, intensity, a_{if}, time, $\mathbf{b}_f$, and gradient, $\mathbf{c}_f$, were subsequently multiplied: $a_{if}\mathbf{b}_f\mathbf{c}_f^T$, as illustrated in Fig. 2. Thus, for each of the seven-

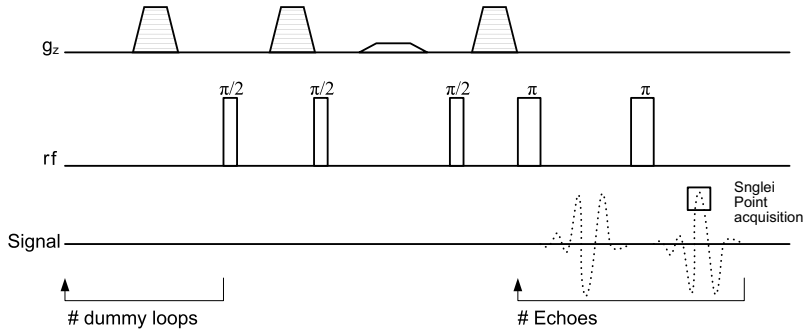


Fig. 4. The pulse program as used to collect the 2D diffusion–relaxation data. Effectively a PGSTE followed by a CPMG echo train, where acquisition is made at peak intensity for every 2nd echo.

teen samples, four 2D-matrices (raw, components 1 and 2 and the PARAFAC-residual) were produced and subsequently imported into Prospa. 2D-Laplace inversions were performed using the optimal regularisation factors, α, as determined by testing a range of reasonable values for each sample, component and residual (Fig. 8). T_2 and D axes, i.e. the observation window, were chosen as follows: $T_{2,\min} = 1 \times 10^{-3}$ s, $T_{2,\max} = 1.5$ s, $D_{\min} = 1 \times 10^{-16}$ m^2 s^{-1} and $D_{\max} = 1 \times 10^{-7}$ m^2 s^{-1} in 24 logarithmically spaced steps in both directions. The T_2 and D boundaries were chosen purposely too wide to include T_2–D space where no intensities should be observed. This helps in determining the quality of fit, keeping the elasticity of the regularisation factor in mind.

T_2–D correlation pairs were identified from the resulting spectra of raw data, components 1 and 2 and the PARAFAC residuals from each sample and quantified by their sum relative to total spectral intensity of each raw sample spectrum, respectively. The qualitative as well as the quantitative aspects were compared for two selected samples and the intensity ratios of components 1 to component 2 were compared to the known ratios of oil and water in the samples.

4. Results and discussion

The major variation introduced in this investigation was imposed by the factorial design varying the major components, starch, gluten and fat of wheat dough, while keeping the water concentration constant. This was, however, not the case of the additional flour samples that included only natural variation and water concentration corresponding to the optimal water absorption (Table 1). Despite the relatively complex composition of dough with protons associated with carbohydrates, protein, fat and water, it was expected that only protons associated with molecules with high rotational and translational (diffusion) freedom could be observed in this experiment. Thus, only water and oil components were expected, possibly influenced by the proportions of other components, i.e. starch and gluten, of the mixture. Wheat flour contains 1.5–2.5% lipids of which roughly one third is mono-, di- and tri-glyceride lipids, one third is starch lipids and one third is bound lipids (glycolipids and phospholipids). In dough, glycolipids and gluten-bound phospholipids are thought to form a laminar phase, stabilising a micro-emulsion of free phospholipids and acyl lipids in water all encapsulated by the gluten network [16].

Table 1 lists the mixture proportions of the samples. The assumed water absorption of 60% in the commercial flour was not correct, as the true value was determined to 64.1%. However, this is of no importance in the current study, since the optimal water absorption of the mixtures was unknown and the water absorption level was set merely to be able to form a visco-elastic dough, regardless of the proportions of starch, gluten, fat and flour. The varying mixing time thus reflects the variation in rheological properties of the various mixtures. As expected, high levels of starch and/or fat prolong the mixing time, while high levels of gluten hardens the dough, resulting in shorter mixing times. The diverse proportions of major ingredients were thus expected to create variation in the micro-environments for water and fat in the mixtures to be explored by the NMR recordings.

The PARAFAC model of the seventeen recorded double exponentially decays was derived from running the model with 1–4 components. The validity of the models was initially explored by simultaneous inspection of the loadings, the explained variance, core consistency and number of iterations [7,17]. From this exercise it was easily concluded that only a model with either two or three components could be valid. The two-component model had a core consistency of 100%, while that of the three-component model was close to 60%. By subsequent inspection of the loadings the third component displayed deviant behaviour from the exponential decay in the gradient mode and thus the two-component model was chosen. The chosen two-component model presented in Fig. 5 explains 99.8% of the variation in the entire data material.

In Fig. 5 the PARAFAC results were normalised to maximum intensity in the exponential decay modes (i.e. modes 2 and 3), thus leaving the relative variation in the sample mode for direct quantification of components (Fig. 5a). Due to the second order advantage of the PARAFAC algorithm the scores in mode 1 need only to be scaled to one reference value in order to give the pure component concentrations. In the two component PARAFAC model, component 1 scores were positively correlated with the added fat content, while component 2 scores were inversely correlated. Since the NMR recordings are only expected to return signals from oil and water, components 1 and 2 were quantitatively assigned to oil and water, respectively, on the basis of the scores (Fig. 5a). In mode 2, the relaxation time mode in Fig. 5b, two distinct exponentially decaying components are resolved; component 1 with a significantly slower decay than component 2, both decaying to approximately zero intensity. In mode 3, the gradient intensity mode in Fig. 5c, two distinct, apparently exponentially decaying components are observed; component 1 has a significantly slower diffusion than component 2. Component 2, however, clearly appears to be multi-exponential with a very slow diffusing covarying feature that does not reach zero intensity within the chosen limits of gradient strength.

A powerful visualisation taking the outer product as illustrated in Fig. 2 of the individual component vectors from mode 2 and mode 3 of the two PARAFAC components is presented in Fig. 5d and e. The combined correlated decays are readily visualised for further interpretation by transformation to the T_2–D domain. It is thus possible by PARAFAC to decompose raw 2D-decays from each sample into their pure components and a sample specific residual matrix. This is done in the left columns of Figs. 6 and 7 for two markedly different samples, S01: S0G0F0 and S08: S2G2F2, presented with equal

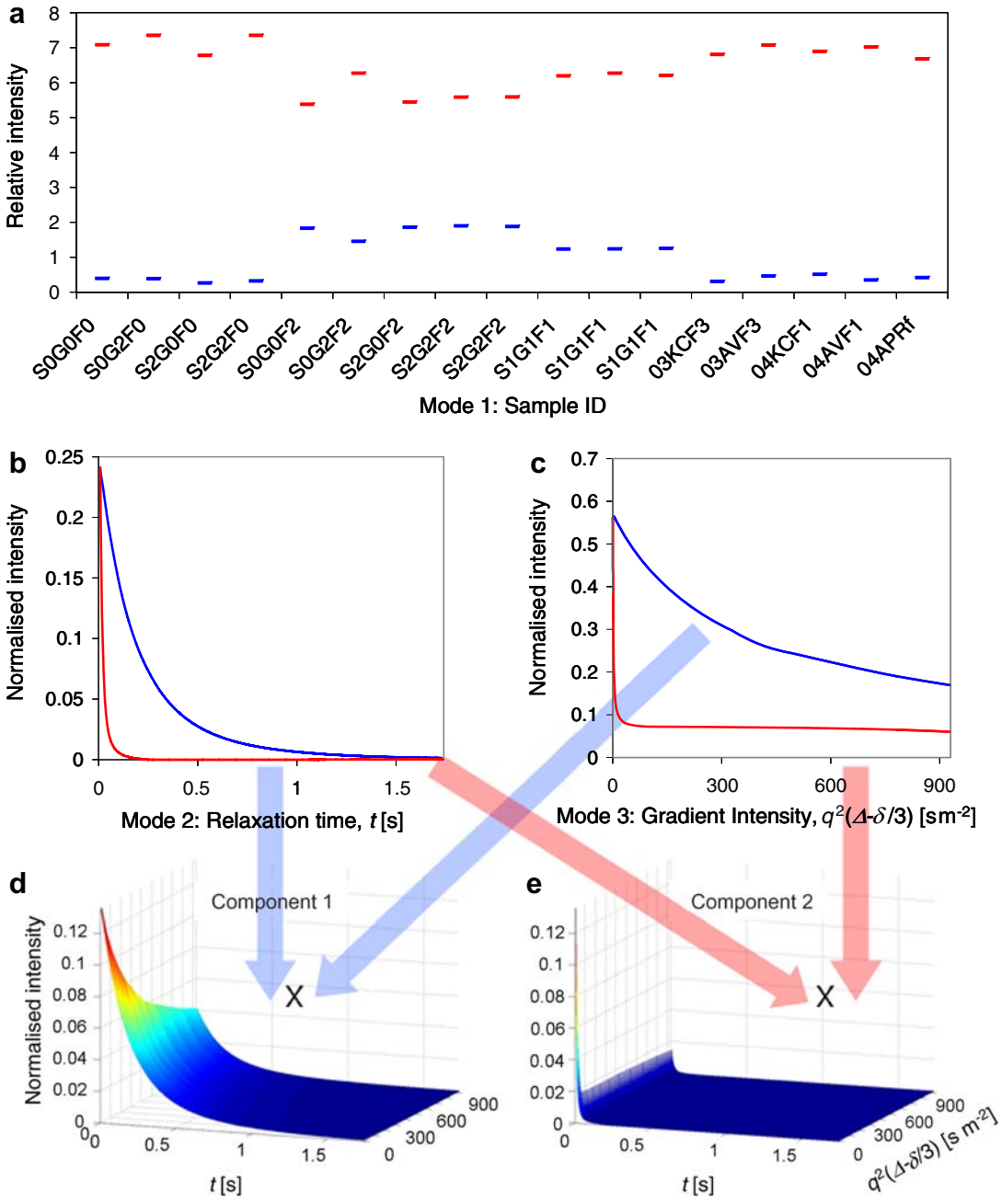


Fig. 5. Two component PARAFAC model of DRCOSY data. — (blue) Component 1, — (red) Component 2. (a) Mode 1: Relative intensities (scores) of components 1 and 2 for each individual sample/mixture. (b) Mode 2: Normalised intensities of components 1 and 2 in the relaxation time direction. (c) Mode 3: Normalised intensities of components 1 and 2 in the gradient direction. (d) and (e) The two DRCOSY components are visualised by taking the outer products of respective components in modes 2 and 3. This figure is equivalent to Fig. 2 explaining the PARAFAC decomposition. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

scaling for proportional interpretation. For interpretation and further quantification the raw data, the two components and the residual 2D-data are transformed into the T_2–D domain by 2D-Laplace inversion in the right column. Figs. 6 and 7 are thus the actual decomposition conceptually presented in Fig. 1.

In Fig. 6 sample S01: S0G0F0 is presented both in the time-gradient domain and in the T_2–D domain. From the raw data (Fig. 6a) it is quite clear that the data is composed of both fast and slow diffusing features as well as fast and slow relaxation features. By the 2D-Laplace inversion (Fig. 6b) this is nicely visualised as nine peaks as T_2–D pairs along with their approximate relative intensities listed in Table 2. Peaks with less than 1% of maximum intensity are not represented in the figure nor in the table, for which reason the sum of listed peaks only sums to 99.1%. In comparison, sample S08: S2G2F2 displays eight peaks in Fig. 7b. Peaks P1–P7 are common for both S01 and S08,

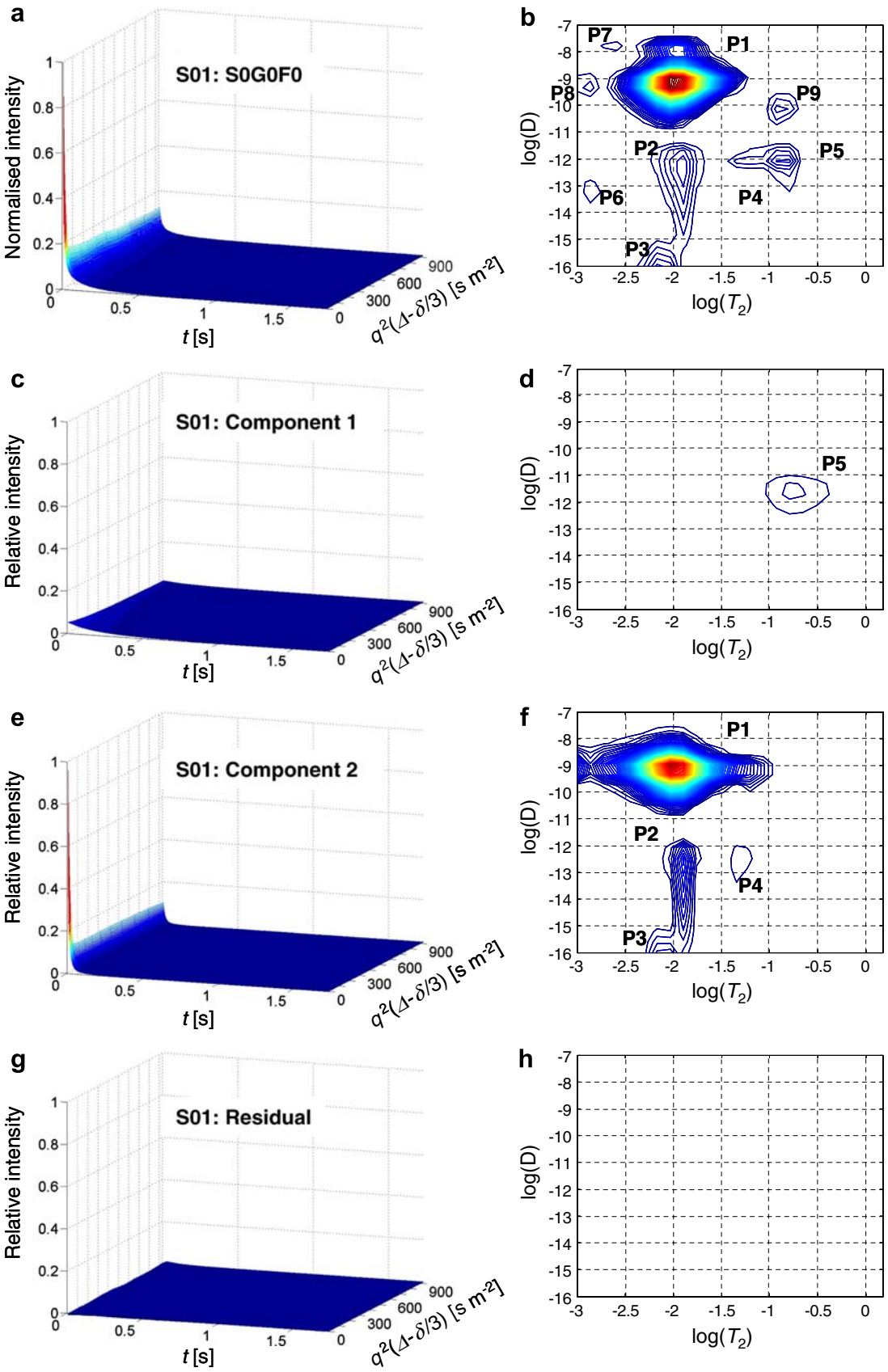


Fig. 6. Sample S01: S0G0F0 split quantitatively into its PARAFAC components 1 + 2 and residual (left column) presented with their corresponding T_2–D correlation maps obtained by 2D-Laplace inversion (right column). Every contour line in the right panels represents additional 1% of maximum intensity.

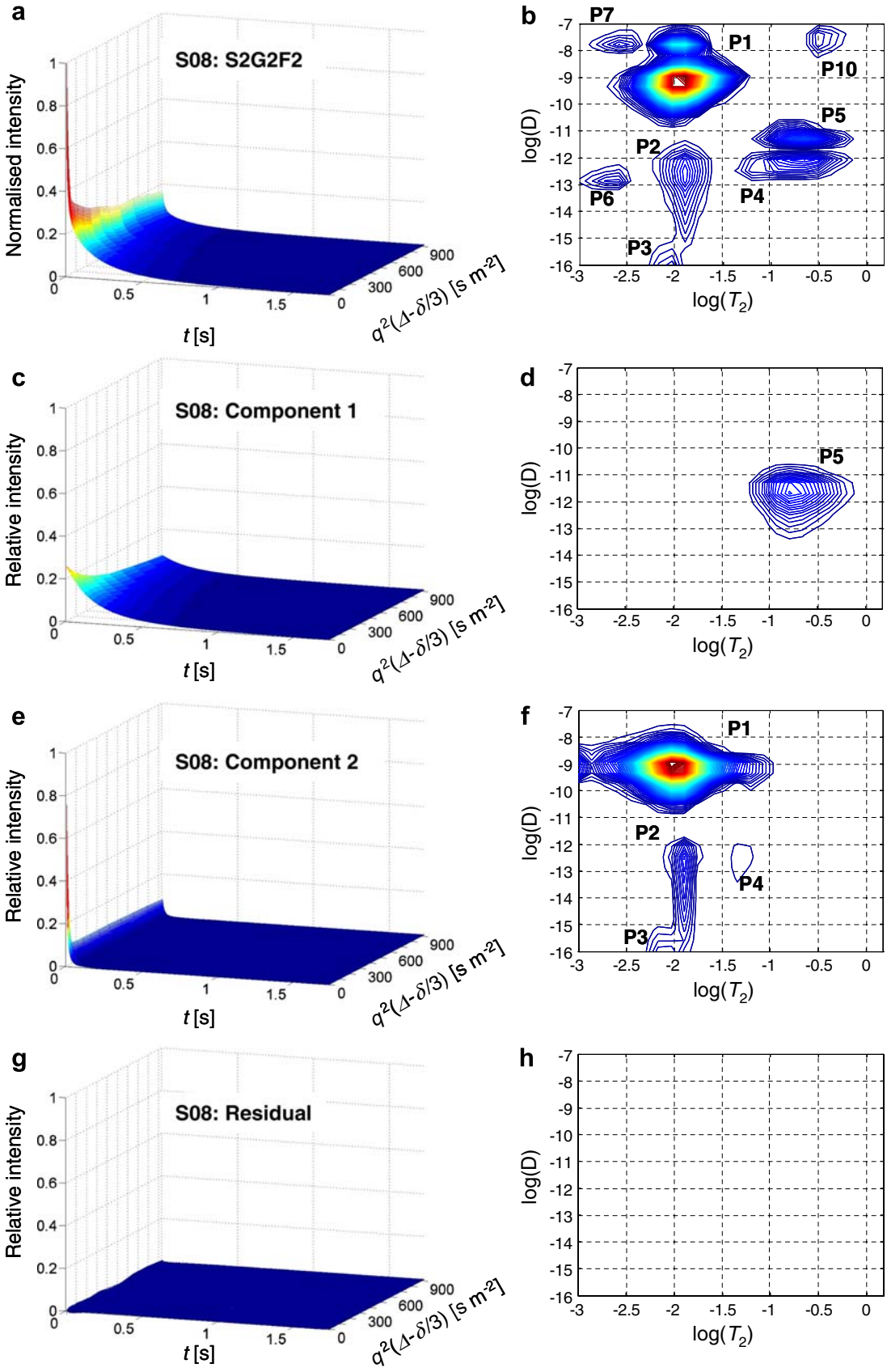


Fig. 7. Sample S08: S2G2F2 split quantitatively into its PARAFAC components 1 + 2 and residual (left column) presented with their corresponding T_2–D correlation maps obtained by 2D-Laplace inversion (right column). Every contour line in the right panels represents additional 1% of maximum intensity.

Table 2
Intensities of peaks P1–P10 as identified in Figs. 6 and 7 in the T_2–D domain for sample S01 and S08

Peak	T_2 [ms]	D[m^2 s^{-1}]	S01: S0G0FO					S08: S2G2F2				
			Raw	PARAFAC				Raw	PARAFAC			
				Comp 1 (%)	Comp 2 (%)	Residual (%)	Sum (%)		Comp 1 (%)	Comp 2 (%)	Residual (%)	Sum (%)
Total int	All	All	100.0	2.2	102.5	0.1	104.8	100.0	12.0	92.8	1.4	106.2
P1	10	1×10^{-9}	88.6		95.9		95.9	78.1		86.8		86.8
P2	13	3×10^{-13}	4.5		4.3		4.3	6.4		3.9		3.9
P3	6	1×10^{-16}	1.5		1.6		1.6	0.9		1.5		1.5
P4	50	1×10^{-12}	0.6		0.6		0.6	0.6		0.6		0.6
P5	180	1×10^{-12}	1.8	2.2			2.2	10.2	12.0			12.0
P6	1.6	5×10^{-14}	0.6				0.0	1.4				0.0
P7	2.5	1.5×10^{-8}	0.2				0.0	1.1				0.0
P8	1.3	5×10^{-10}	0.5				0.0	0.0				0.0
P9	130	8×10^{-11}	0.8				0.0	0.1				0.0
P10	300	3×10^{-8}	0.0				0.0	1.1				0.0
Sum			99.1	2.2	102.5	0.0	104.6	99.6	12.0	92.7	0.0	104.8
Residual			0.9	0.0	0.0	0.1	0.1	0.4	0.0	0.0	1.4	1.4

Intensities are presented as relative to total intensity of the respective raw spectrum. Peaks are summed leaving an unrealised residual for each spectrum, i.e. Raw, Comp 1, Comp 2 and PARAFAC-Residual. The PARAFAC-Laplace peaks are summed horizontally.

while P8 and P9 are only present in S01 and P10 is present only in S08.

The challenges in the interpretation of the similarities and deviations of these and the remaining samples are overwhelming, although some delimitation can be put forward by manual inspections. Peaks with diffusion coefficient higher than say 10^{-8} m^2 s^{-1} must be considered noise, since the limit of diffusion in this system must be considered to be the diffusion of free water, i.e. 2×10^{-9} m^2 s^{-1}. Likewise, peaks below $T_2 = 10$ ms are unlikely, since the duration of the pulsed gradient time (7 ms) and the storage time (16.5 ms) would probably render such signals extinct prior to the first recordings by the CPMG train. Thus, P6, P7, P8 and P10 in Figs. 6b and 7b are artefacts due to noise and the ill-conditioned Laplace inversion. P3 is a typical border phenomenon due to a small offset in the raw data along the gradient direction. Thus, by reasoning it is possible to reduce the problem to five peaks of interest. Apart from peak P9 present only in S01, all relevant peaks are present in both samples. P1 is by far the most intense with 88.6% for S01 and 78.1% for S08 (Table 2) and considering the high diffusion coefficient it must be water being able to move freely in the matrix, although restricted in rotational freedom ($T_2 = 10$ ms), thus probably associated with the surfaces of the matrix, i.e. the gluten network. P2 at 4.5% and 6.4%, respectively, equally restricted in rotational freedom is much more restricted in diffusion with a broad distribution of diffusion coefficients that indicate tightly bound water, probably associated with water absorption by swelling starch granules. P4 is very small, 0.6% for both, and only identified as a peak due to the new method (see Section 4 below). P5 varies markedly between the two samples, 1.8% in S01 and 10.2% in S08, proportional with the difference in oil content for the two samples. In Fig. 7b, S08 peak P5 forms a double peak which will be discussed below. With $D = 10^{-12}$ and $T_2 = 180$ ms P5 is most probably fat restricted in diffusion by the size of the vesicles in the water matrix and at the same time with more slow relaxation than the "free" water molecules.

4.1. Qualitative analysis

We will now demonstrate the experiment interpretation from the point of view of the new method. The samples S01 and S08 were decomposed into the two PARAFAC components (Figs. 6c, e and 7c, e) corresponding to their relative amounts in the raw data and the sample-specific non-systematic residual (Figs. 6g and 7g). The residual is the difference between original sample recording and the PARAFAC components. While transforming the PARAFAC components 1 and 2 and the residual rather than the raw data by 2D-Laplace inversion only common structures to the entire set of samples are investigated (Figs. 6d, f, h and 7d, f, h). Keeping in mind that 99.8% of all the variation in the set of 2D-multi-exponential data is explained by just two PARAFAC components justifies this approach when investigating these data.

First we observe that only peaks P1–P5 are represented in the PARAFAC-Laplace components 1 (Fig. 6d, f) and 2 (Fig. 7d, f). The residuals did not contain any significant exponential behaviour, which was also the case for all other samples and accordingly the T_2–D plots are empty (Figs. 6h and 7h). Component 1 contains one peak only (Figs. 6d and 7d)—the P5 fat component described above. Being a PARAFAC component shows that this component varies between samples independently from the other components. This is in fine accordance with the fact that oil content was varied purposely in the design of the experiment. PARAFAC-Laplace component 2 in Figs. 6f and 7f contains P1–P4, in which P4 was only recognised in its own right by this method. In the T_2–D spectra of the raw data this was not readily observed, as this peak was both small and overlapping with P5. The PARAFAC-Laplace decom-

position thus shows that the water found in three different compartments in the matrix is highly covarying, which in turn indicates that the water compartmentalisation in this study of dough at 34 °C is highly unaffected by the proportions of starch, gluten and oil. In Table 2 the proportionate intensities of the peaks in the PARAFAC-Laplace components are presented relative to total intensity of the summed peaks of the *raw* Laplace data. Due to the non-uniqueness and labile character of the Laplace algorithm these intensities do not add up to 100%—but relatively close. The relative proportions of the components (Component 1/Component 2) in S01, 2.2%/102.5% directly correspond to the summed proportions, 1.8%/95.2% in the spectrum of the raw data as in the S08 case where the 12.0%/92.7% component ratio corresponds to the 10.2%/86.0% ratio of the summed raw data. This quantitatively confirms that the PARAFAC-Laplace decomposition is equivalent to ordinary direct 2D-Laplace inversion. The relative quantitative information is conserved.

The fact that the two components when transformed into the T_2–D domain have only a few distinct features suggests that other features found (i.e. P6–P10) in the individual raw spectra are only artefacts inflated by the ill-conditioned properties of the Laplace inversion. The Laplace inversion estimates many parameters based on one sample only which gives a poor independent variable to parameter ratio. Although every 2D-landscape is generated from multiple scans of thousands of data points, they are strongly covarying, thus giving a single error the possibility to be inflated by the data analysis. Artefacts from unfortunate sample presentation in the spectrometer, such as different packing of the material, air bubbles or bad mixing of the sample may turn up in the spectra and mistakenly be interpreted as components. PARAFAC resolvation prior to 2D-Laplace inversion significantly reduces these artefacts by the inherent second order advantage of reducing noise and uniquely extracting pure covarying components.

The optimal regularisation factor in the 2D-Laplace inversion was calculated independently for raw sample spectra, PARAFAC components 1 and 2 and PARAFAC residuals individually. A high factor produces results with many sharp features, while a low factor produces a smooth result. The optimum is somewhere in between, where the residual variance expressed by χ^2 is close to the mathematical optimum, i.e. lowest obtainable value that at the same time produces relatively smooth and interpretable plots. The actual regularisation chosen is somewhat subjective and based on experience. From Fig. 8 the regularisation factors were chosen to the nearest order of magnitude. Note that χ^2 for the PARAFAC residuals did not vary significantly in the α-plot, as it is noise being modelled by the 2D-Laplace inversion algorithm with no or insignificant exponential behaviour left. In practice, the regularisation factor is quite difficult to determine from sample to sample, but should ideally be the same for all samples in order to interpret and compare spectra, i.e. with the same smoothness independent of the variations in signal-to-noise ratio from sample to sample.

The signal/noise ratios for the PARAFAC components do not vary from sample to sample, as they are represented in all samples—just in different proportions. This leads to a key point that only these two—not seventeen α's need to be determined, i.e. $\alpha_{opt} = 10^8$ for component 1 and $\alpha_{opt} = 10^9$ for component 2 (Fig. 8). The fact that components 1 and 2 use different regularisation factors is due to their difference in intensity and thus different signal/noise ratio. In this example the most intense component dominates when determining the regularisation of the raw sample data. This is nicely illustrated by the samples presented in Figs. 6b and 7b in which peak, P5, appears as a single peak in T_2–D spectrum of S01, while it is a double peak in S08. The water peak, P1, is the dominant signal and will as an apparent optimum choose $\alpha_{raw} = 10^9$ (Fig. 8), even though this value is not optimal in relation to the weaker peak(s). Thus, non-significant features and noise in the data lead to confusing sharp peaks, double peaks (Fig. 7b) and non-physical peaks in spectra. On the other hand, focusing on the weaker peaks in the raw data by choosing a smaller α would eliminate non-physical peaks, but at the same time broaden the more intense peak(s) and even intercept smaller true peaks. PARAFAC analysis prior to the 2D-Laplace inversion elegantly solves this dilemma by separating the true physico-chemical components prior to the change of domain by Laplace inversion.

Because artefacts are introduced by the Laplace transformation as described above and because the choice of regularisation factor plays a significant role depending on the signal/noise ratios of the sample spectra the theoretically possible route of running PARAFAC in the T_2–D domain as indicated in Fig. 1 by arrow with a question mark was quickly abandoned.

4.2. Quantitative analysis

The ratio between oil and water can be studied quantitatively, and these data are presented in Fig. 9. A theoretical ratio is calculated from a commonly accepted level of acyl lipids in flour (0.7%) and the amount of fat which was added to the samples. In both water and oil the approximate number of H-atoms relative to molecular mass is 1/9, thus the relative abundance of fat–H to water–H can be calculated directly by their masses and by their presence in the DRCOSY data. Component 1 versus component 2 ratios are plotted for each sample directly from the PARAFAC scores in mode 1 (Fig. 5a) as well as from PARAFAC-Laplace spectra of the same components and summed to total intensity as done for S01 and S08 in Table 2.

Fig. 9 shows the quantitative ratio determined by the PARAFAC scores as well as PARAFAC-Laplace volumes of the components 1 and 2. They are both perfectly correlated ($r = 0.971$) with calculated oil/water ratios in the samples. In the calculation of the natural oil content the

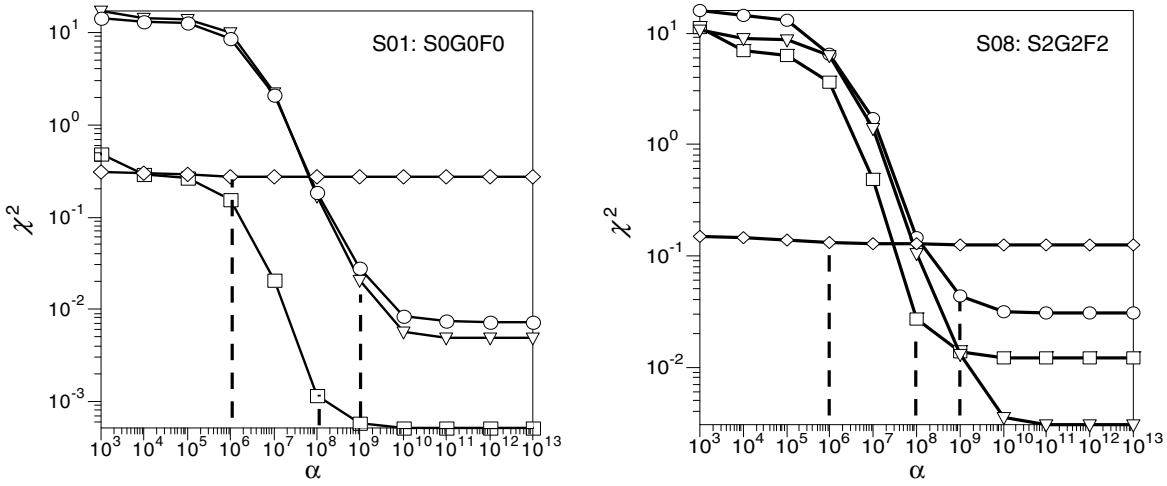


Fig. 8. Residual variance, χ^2 as a function of regularisation factor, α for sample S01: S0G0F0 and S08: S2G2F2: Raw data (○) $\alpha_{opt} = 10^9$, PARAFAC-component 1 (□) $\alpha_{opt} = 10^8$, PARAFAC-component 2 (▽) $\alpha_{opt} = 10^9$ and the PARAFAC-residual (◇) $\alpha_{opt} = 10^6$.

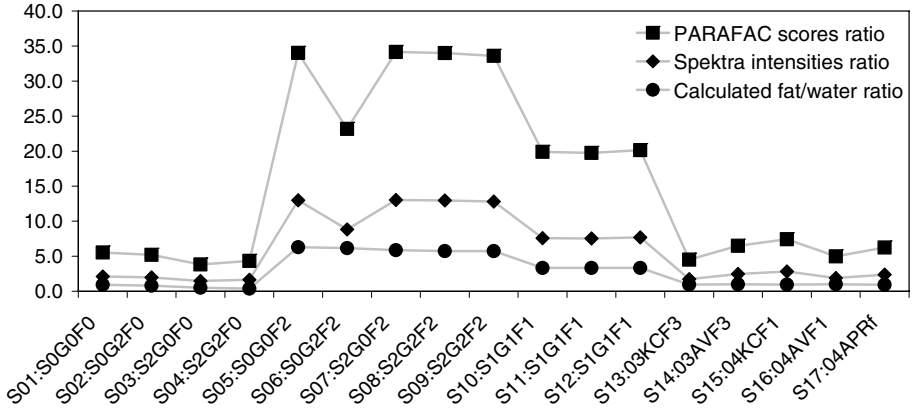


Fig. 9. Calculated water/oil ratios plotted with PARAFAC scores ratios and PARAFAC-Laplace volume ratios of component 1/component 2.

average oil contribution from flour (0.7%) was included as an offset. However, the gluten added in the mixtures may also contain significant amounts of lipids, as these are usually not easily extracted from gluten without destroying the gluten [16]. However, as no quantitative information was available, the eventual gluten bound fat was neglected in the calculations. The PARAFAC scores ratio and the summed PARAFAC-Laplace volume ratios of the two components are perfectly correlated, i.e. $r = 1.000$, because the 2D-Laplace inversion is performed on the exact same two components for each sample. Thus, for pure quantification purposes the Laplace domain transformation may actually be superfluous.

The PARAFAC scores ratios in Fig. 9 do not overlap the calculated fat/water ratios, because the PARAFAC analysis cannot take the relaxation during gradient encoding in the first 7.8416 ms into account. Fast relaxing compounds (say, $T_2 < 20$ ms) are thus significantly underestimated. The ratios of the summed T_2–D spectra of the two components should ideally be exactly overlapping the calculated ratios. However, some of the water signal was probably lost during acquisition due to T_2 and T_1 relaxation during gradient and storage time.

In Fig. 9 sample S06: S0G2F2 indeed looks like an outlier (knowing the fat content to be high). This is probably due to bad mixing or packing of the material in the tube prior to the recording. However, this outlying sample does not destroy the PARAFAC-Laplace model—it is only the proportions of water and oil that seem unlikely, taking prior knowledge into account. When leaving this sample out of the model, the correlation coefficient between the calculated and the PARAFAC estimated fat/water ratios was: $r = 0.997$.

Although both the PARAFAC ratios and the spectra ratios of the components are far from the known fat/water content, the near 100% correlation with known fat/water ratio is useful. Knowing the actual fat/water ratio of just one sample, the remaining samples can be calculated from the PARAFAC scores ratios, even if the fast relaxing water relaxation signal was not recorded quantitatively. That is the 2nd order advantage which is of great value in cases where the method of recording the signal, i.e. the

DRCOSY, cannot be set optimally for quantitative recording of the signals from fast relaxing components. As long as all 2D diffusion–relaxation spectra are recorded identically, relative comparisons are always possible based on the PARAFAC scores only. The unique PARAFAC resolvation of all varying components is based on having 2D-matrices of data for each sample, rather than just 1D-vectors.

5. Conclusions

The method combining the unsupervised PARAFAC model with 2D Laplace inversion has shown to be a significant improvement in the analysis of two-dimensional multi-exponential data recorded by DRCOSY. It allows identification and quantification of pure components and it reduces artefacts and stabilises the subsequent 2D Laplace inversion. This approach supports research by identifying real systematically varying components while filtering artefacts associated with unfavourable conditioning of the Laplace inversion.

The new procedure can be regarded as a step towards automatic analysis of DRCOSY and similar data as an alternative to biased human interpretation. The analysis requires a homologous set of samples in which specific factors of interest have been varied either by experimental design or by natural diversity of the materials investigated. In a future publication we will demonstrate that double SLICING [18–20] can be used to extract discrete T_2–D components from a single COSY recording.

Acknowledgments

A number of institutions, private companies and individuals are greatly acknowledged: The MacDiarmid Institute at Victoria University Wellington, New Zealand for providing time, space and knowledge; Allan K. Hardacre, Crop & Food Research, Palmerston North, New Zealand for good discussions and providing equipment. Bala Diagaradjan, Weston Milling, Lower Hutt, Wellington, New Zealand for reference analysis on short notice; Cerealia Mills, Vejle, Denmark for facilitating laboratory space and equipment in the milling process. For financial support we are greatly thankful to DFFE with the project title ''Un-entangling complex food systems by NMR/US spectroscopy and mathematical modelling'' sponsored by the Ministry of Food, Agriculture and Fisheries, to the Interreg IIIA programme between Fyns Amt, Odense, Denmark and Technologie-Region K.E.R.N, Rendsburg, Germany and to University of Copenhagen, Faculty of Life Sciences, Frederiksberg, Denmark for financing first author's Ph.D scholarship. Special thanks to Q-Interline Spectroscopic Analytical Solutions, Roskilde, Denmark, Knud Højgaards Fond, Charlottenlund, Denmark and Danmarks Jordbrugsvidenskabelige Ph.D-forening, Copenhagen, Denmark for travel Grants making the visit to Wellington possible for the first author.

References

[1] P.T. Callaghan, S. Godefroy, B.N. Ryland, Use of the second dimension in PGSE NMR studies of porous media, Magn. Reson. Imaging 21 (2003) 243–248.
[2] P.L. Hubbard, P.J. Watkinson, L.K. Creamer, A. Gottwald, P.T. Callaghan, Two-dimensional Laplace inversion NMR technique applied to the molecular properties of water in dry-salted mozzarella-type cheeses with various salt concentrations, in: S.B. Engelsen, P.S. Belton, H.J. Jakobsen (Eds.), Magnetic resonance in food science. The multivariate challenge. The proceedings of the 7th international conference on applications of magnetic resonance in food science held in Copenhagen on the 13–15th September 2004, The Royal Society of Chemistry, Cambridge, UK, 2005, pp. 225–232.
[3] Y. Qiao, P. Galvosas, P.T. Callaghan, Diffusion correlation NMR spectroscopy study of anisotropic diffusion of water in plant tissues, Biophys. J. 89 (4) (2005) 2899–2905.
[4] J. Christensen, L. Nørgaard, R. Bro, S.B. Engelsen, Multivariate autofluorescence of intact food systems, Chem. Rev. 106 (6) (2006) 1979–1994.
[5] R. Bro, P.I. Hansen, N. Viereck, M. Dyrby, H.T. Pedersen, S.B. Engelsen, A new principle for unique spectral decomposition of 2D NMR data, in: S.B. Engelsen, P.S. Belton, H.J. Jakobsen (Eds.), Magnetic resonance in food science. The multivariate challenge. The proceedings of the 7th international conference on applications of magnetic resonance in food science held in Copenhagen on the 13–15th September 2004, The Royal Society of Chemistry, Cambridge, UK, 2005, pp. 195–203.
[6] H.T. Pedersen, M. Dyrby, S.B. Engelsen, R. Bro, Application of multi-way analysis to 2D NMR data, Annu. Rep. NMR Spectrosc. 59 (2006) 207–233.
[7] R. Bro, PARAFAC. Tutorial and applications, Chemom. Intell. Lab. Syst. 38 (2) (1997) 149–171.
[8] P.T. Callaghan, Principles of Nuclear Magnetic Resonance Microscopy, Clarendon Press, Oxford, 1991.
[9] E.L. Hahn, Spin echoes, Phys. Rev. 80 (4) (1950) 580–594.
[10] J.E. Tanner, Use of the stimulated echo in NMR diffusion studies, J. Chem. Phys. 52 (1970) 2523–2526.
[11] H.Y. Carr, E.M. Purcell, Effects of diffusion on free precession in nuclear magnetic resonance experiments, Phys. Rev. 94 (1954) 630–638.
[12] S. Meiboom, D. Gill, Modified spin-echo method for measuring nuclear relaxation times, Rev. Sci. Instrum. 29 (1959) 688–691.
[13] M.D. Hurlimann, L. Venkataramanan, C. Flaum, The diffusion-spin relaxation time distribution function as an experimental probe to characterize fluid mixtures in porous media, J. Chem. Phys. 117 (2002) 10223–10232.
[14] Y.-Q. Song, L. Venkataramanan, M.D. Hürlimann, M. Flaum, P. Frulla, C. Straley, T_1–T_2 correlation spectra obtained using fast two-dimensional laplace inversion, J. Magn. Res. 154 (2002) 261–268.
[15] C.A. Andersson, R. Bro, The *N*-way Toolbox for MATLAB. Chemom. Intell. Lab. Syst. 52(1) (2000) 1–4, <http://www.models.life.ku.dk/source/nwaytoolbox/>.
[16] H.-D. Belitz, W. Grosch, Food Chemistry, second ed., Springer, Verlag, Berlin, 1999.
[17] C.A. Andersen, R. Bro, Practical aspects of PARAFAC modelling of fluorescence excitation-emission data, J. Chemom. 17 (4) (2003) 200–215.
[18] H.T. Pedersen, R. Bro, S.B. Engelsen, Towards rapid and unique curve resolution of low-field NMR relaxation data: trilinear SLICING versus two-dimensional curve fitting, J. Magn. Res. 157 (2002) 141–155.
[19] S.B. Engelsen, R. Bro, PowerSlicing, J. Magn. Res. 163 (2003) 192–197.
[20] C. Manetti, C. Castro, J.P. Zbilut, Application of trilinear SLICING to analyse a single relaxation curve, J. Magn. Res. 168 (2004) 273–277.