Low-field Nuclear Magnetic Resonance and Chemometrics Applied in Food Science

Centre for Critical Quality Attribute Determination in Muscle Foods

PhD Dissertation by

Henrik Toft Pedersen

Supervisor

Søren Balling Engelsen

Food Technology

Department of Dairy and Food Science

The Royal Veterinary and Agricultural University

Preface

The work performed during the three-year period of my PhD is part of a research project called “Centre for Critical Quality Attribute Determination in Muscle Foods”. My work has been centred on development of new low-field NMR as well as chemometric techniques with the aim to improving performance and applicability in food analysis.

Apart from a six-month stay at Unilever (Colworth House, UK), my work has mainly been performed at the Food Technology group, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University.

I am thankful to the people of the Food Technology group for all the help and valuable discussions over the years, in particular Rasmus Bro, Claus Andersson and Søren B. Engelsen from whom I received help with mathematical and algorithmic problems. I also want to thank Lars Nørgaard for filling some of the gaps in my knowledge about the science of chemometrics and professor Lars Munck for splendid enthusiasm. Last but not least, the other PhD students in the group should be mentioned.

Steve Ablett from Unilever receives my gratitude for allowing me to come and work in his lab.

Tim Benson and Thierry Guiheneuf from Resonance Instruments should also be mentioned for helping me off to a good start with the low-field NMR instrument as well as valuable support later, when problems or questions arose that I could not solve myself.

In addition, the Danish Research Council (SJVF) is thanked for the financial support, which made this work possible.

11 December 2001

Summary

Since the introduction of low-field nuclear magnetic resonance (LF-NMR) instruments in food science a tremendous development has taken place regarding the capabilities of the hardware applied as well as the performance of data analytical techniques available.

This thesis primarily deals with the progress of the analysis of LF-NMR data as well as the techniques that are commonly used, and some future possibilities for further improvements are suggested.

Focus is on papers published combining the three cornerstones of the research performed in connection with this thesis: LF-NMR, food materials and chemometrics.

It is clear from the work performed that LF-NMR has great potential within food science, since LF-NMR measurements are performed on a volume of sample rather than just on the surface. This can compensate for the inhomogeneity of most food samples. Furthermore, little or no sample preparation is generally required and no chemicals needed. The fact that all protons in a sample contributed to the LF-NMR signal, the state and distribution of water in samples can be probed, diffusion constants can be measured and it is possible to measure frozen samples or freeze or cook the sample inside the probe-head while continually measuring.

The instrument allows application of a variety of different pulse experiments each probing different proton populations in the sample. It is therefore obvious that there is a large range of possibilities for enhancing the information content of the acquired data manipulating the sample for enhanced analysis. Many of these possibilities are not available in other spectroscopic techniques.

The papers that constitute my own contribution to the research in this area can be divided into two areas: application and data analysis. The papers are concerned with the topic of combining LF-NMR on food materials and chemometrics [P1,P3,P4,P6,P8] as well as development of new methods both in LF-NMR and chemometrics [P2,P5,P6,P7].

It is shown that in most cases the multivariate data analysis offers great potential for improving the traditional analysis of water and fat in fish flesh [P1], oil, water and protein content in rape and mustard seeds [P3], staling and baking of bread [P4] as well as water and fat content of meat [P6]. In addition new possibilities in data acquisition are proposed in terms of extending pulse experiments towards acquisition of data structures designed for multivariate rather than univariate data analysis [P6] and new designs of NMR hardware are tested for analysis of food-related materials [P8].

Existing multivariate methods are compared [P2] and algorithms proposed for different kinds of data are altered to handle LF-NMR relaxation data with the purpose of enhancing explorative as well as predictive data analysis [P5,P7].

Sammendrag

Siden indførelsen af lavfelts kærnemagnetisk resonans (LF-NMR) instrumenter i levnedsmiddelforskningen er der sket en voldsom udvikling, både hvad angår hardware egenskaber og dataanalysemuligheder.

Denne afhandling handler om udviklingen af LF-NMR dataanalyse og LF-NMR metoder, men traditionelle metoder, som hyppigt anvendes inden for begge områder, vil også blive gennemgået i denne sammenhæng.

Fokus er på publicerede artikler der er centreret om de tre hjørnesten, som danner baggrund for forskningen rapporteret i denne afhandling: LF-NMR, levnedsmidler og kemometri.

På baggrund af det arbejde som er foretaget synes det klart, at LF-NMR har et stort potentiale inden for levnedsmiddelforskningen, eftersom LF-NMR måler på et volumen i stedet for kun på prøveoverfladen. Dette er vigtigt for at kompensere for den kompleksitet og inhomogenitet, som karakteriserer de fleste levnedsmidler. Som en yderligere fordel kræver LF-NMR ingen prøveforberedelse og involverer ikke brug af kemikalier.

Alle protoner i en prøve bidrager til det registrerede signal og man kan derfor måle tilstanden og fordelingen af vand i en prøve, man kan måle diffusionskonstanter, og det er muligt at måle på frosne prøver, samt at fryse eller koge prøver inde i målecellen samtidigt med at man måler. Yderligere tillader instrumentet, at man kan anvende forskellige pulseksperimenter, som hver især giver information om forskellige egenskaber. Det er således klart, at der findes en lang række muligheder for a manipulere protonerne i prøven for at forbedre analysen. Mange af disse muligheder findes ikke for andre spektroskopiske analysemetoder, og dette gør LF-NMR, i kombination med kemometri, til en meget alsidig metode med brede anvendelsesområder inden for levnedsmiddelforskningen såvel som andre områder. De artikler, som udgør mit eget bidrag til forskningen inden for disse områder, kan deles ind i to kategorier: applikationer og dataanalyse. Artiklerne omhandler primært LF-NMR i kombination med kemometri anvendt på levnedsmidler [P1,P3,P4,P6,P8], men udvikling af nye metoder inden for både LF-NMR og kemometri er også behandlet [P2,P5,P6,P7].

Det er blevet vist, at multivariate dataanalysemetoder i de fleste tilfælde tilbyder store muligheder for forbedringer af den traditionelle analyse af vand og fedt i fisk [P1], for vand, olie og protein i raps- og sennepsfrø [P3], for monitorering af staling og bagning af brød [P4] og for analyse af vand- og fedtindhold i kød [P6]. Nye muligheder er også blevet foreslået i form af omskrivning af eksisterende pulseksperimenter, således at multivariate data i stedet for univariate data bliver opsamplet med efterfølgende multivariat dataanalyse [P6], og nyt hardware layout er blevet testet til analyse af levnedsmiddelrelaterede prøver [P8].

Eksisterende multivariate metoder er blevet sammenlignet [P2], og algoritmer som oprindeligt er foreslået til lidt andre typer data, er blevet modificerede og optimerede til at håndtere LF-NMR data. Formålet har været at forbedre eksplorative såvel som prediktive multivariate dataanalyser [P5,P7].

List of papers

P1 Jepsen, S.M.; Pedersen, H.T.; Engelsen, S.B.: Application of chemometrics to low-field ^{1}H NMR relaxation data of intact fish flesh, *Journal of the Science of Food and Agriculture*, 1999, **79**, 1793-1802

P2 Bechmann, I.E.; Pedersen, H.T.; Nørgaard, L.; Engelsen, S.B.: *Comparative chemometric analysis of transverse low-field ^{1}H NMR relaxation data*, in "Advances in Magnetic Resonance in Food Science" (Belton, P.S.; Hills, B.P.; Webb, G.A., eds.), pp. 217-225, The Royal Society of Chemistry, Cambridge, 1999

P3 Pedersen, H.T.; Munck, L.; Engelsen, S.B.: Low-field ^{1}H nuclear magnetic resonance and chemometrics combined for simultaneous determination of water, oil, and protein contents in oilseeds, *Journal of the American Oil Chemists' Society*, 2000, **77** (10), 1069-1076

P4 Engelsen, S.B.; Jensen, M.K.; Pedersen, H.T.; Nørgaard, L.; Munck, L.: NMR-baking and multivariate prediction of instrumental texture parameters in bread, *Journal of Cereal Science*, 2001, **33** (1), 59-69

P5 Pedersen, H.T.; Bro, R.; Engelsen, S.B.: *SLICING - A novel approach for unique deconvolution of NMR relaxation decays*, in "Magnetic Resonance in Food Science: A view to the Future" (Webb, G.A.; Belton, P.S.; Gill, A.M.; Delgadillo, I., eds.), pp. 202-209, The Royal Society of Chemistry, Cambridge, 2001

P6 Pedersen, H.T.; Berg, H.; Lundby, F.; Engelsen, S.B.: The multivariate advantage in fat determination in meat by bench-top NMR, *Innovative Food Science and Emerging Technologies*, 2001, **2**, 87-94

P7 Pedersen, H.T.; Bro, R.; Engelsen, S.B.: Rapid and unique curve resolution of low-field NMR T_2-components by SLICING and trilinear decomposition, *In prep.*, 2001

P8 Pedersen, H.T.; Ablett, S.; Martin, D.R.; Mallett, M.J.D.: Application of the NMR-MOUSE to food emulsions, *In prep.*, 2001

Abbreviations

2D	two dimensional – a matrix
3D	three dimensional – a cuboid
ANOVA	analysis of variance
a.u.	arbitrary unit
CPMG	Carr-Purcell-Meiboom-Gill
DECRA	direct exponential curve resolution algorithm
DTLD	direct trilinear decomposition
FID	free induction decay
GRAM	generalised rank annihilation algorithm
INVREC	inversion recovery
LF-NMR	low-field nuclear magnetic resonance
MLR	multiple linear regression
MOUSE	mobile universal surface explorer
MRI	magnetic resonance imaging
PARAFAC	parallel factor analysis
PCA	principal component analysis
PCR	principal component regression
PFG	pulsed field gradient
PLS	partial least squares
RF	radio frequency
RMSECV	root mean square error of cross validation
RMSEP	root mean square error of prediction
SSE	square sum of error
T	Tesla
T_1	longitudinal or spin-lattice relaxation time constant
T_2	transverse or spin-spin relaxation time constant
WHC	water holding capacity

Table of contents

1 INTRODUCTION

1.1 PRESENTATION OF THE THESIS

It is the aim of this doctoral thesis to give the reader insight into the work that has previously been performed using low-field nuclear magnetic resonance (LF-NMR) in food science in general and to present the work performed during my PhD study.

The presentation will be divided into three main parts: Chapter 2 describing NMR history and theory, including pulse experiments; Chapter 3 dealing with data analysis starting with the traditional methods and procedures, followed by a description of the work published where LF-NMR and chemometrics have been combined; Chapter 4 in which some ongoing and possible future work on the application of LF-NMR in food science is described and suggested.

1.2 DESCRIPTION OF PROJECT

In 1998 a collaboration between Food Technology and Meat Science at The Royal Veterinary and Agricultural University (KVL), the Department of Product Quality at the Danish Institute of Agricultural Sciences (DJF), the Department of Seafood Research at Danish Institute for Fisheries Research (DFU) and the Magnetic Resonance Research Centre at Århus University (located at Skejby Hospital) was started on the basis of a grant from The Danish Research Council (SJVF). In the application to SJVF, the following summary was included to describe the purpose of the collaboration:

> "This project is designed to provide a high-powered research centre between leading agencies working on the properties of meat and fish. It will be a platform on which other initiatives are built and which will provide major insights into basic issues in muscle food production technology and quality attributes.
>
> Attributes such as the water-holding capacity (WHC) of muscle foods are of major concern for the meat and fish industry, as they affect critical technological traits (processing yield), important sensory characteristics (appearance, flavour, tenderness, juiciness), nutritional characteristics (water soluble vitamins, minerals) and

processing characteristics (stickiness, salt uptake). Low WHC is most undesirable, resulting in unacceptably high drip loss, which may occur at many stages during processing. The industry is therefore vitally interested in factors controlling this problem.

The ability to use rapid, continuous, non-destructive and non-invasive techniques to measure WHC would provide the industry with means of selection and control of product in a manner not possible at present with consequent better use of raw materials, improved yields, consistency of product characteristics and hence better financial returns. Current methods do not achieve these objectives.

The investigations in this project will take a novel approach on well-described materials to find solutions to these pressing problems. The use, at the highest level, of sophisticated spectroscopic techniques, such as NMR (Nuclear Magnetic Resonance) and fluorescence, with the most recent instruments and with new probes to be developed, will be coupled with the ability to analyse the complex data obtained by application and development of the most advanced chemometric techniques in a manner not employed for this task previously.

In addition, these results will be linked with pre-mortem and immediate post-mortem energy metabolism and with detailed structural, biochemical and physical measurements on the same, or comparable, materials to provide new insights into controlling the processes of change in muscle food quality attributes.

To do this the project brings together considerable resources, expertise and knowledge of pork muscle, fish muscle, measurement techniques and chemometrics from leading university and sector research institutes in the food and medical fields.“

The work presented here has been carried out as part of this project and due to the different expertise and facilities at the institutions involved, emphasis in this work has been given to the exploration of new possibilities in LF-NMR and chemometrics in food science. The aim has been to test existing protocols

and if possible to develop new protocols to produce data with more information specific to the problem in question and further to develop and test new chemometric possibilities in combination with LF-NMR relaxation data to enhance performance of the data analysis.

2 NUCLEAR MAGNETIC RESONANCE

2.1 HISTORY

The theory of nuclear magnetic resonance (NMR) was put forth in the mid-1930's by the Dutch physicist Gorter [1], but it was not until 1945 that the American physicists Bloch and Purcell discovered NMR in the form that is known today [2,3]. The possibilities of NMR in chemistry were not appreciated until 1950, when Proctor and Yu discovered the chemical shift which enabled elucidation of molecular structures of organic compounds [4]. Although originating in physics, it is in chemistry that NMR has provided the greatest possibilities.

In the first continuous wave NMR instruments spectral resolution was poor, but with the introduction of signal averaging, sensitivity improved significantly, and by introducing pulsed NMR instruments along with Fourier transformation of the acquired time-domain relaxation, signal resolution was even further improved. By introducing pulsed NMR the time of acquisition was reduced as an additional advantage. It was discovered that, particularly for solids, spinning of the sample inside the instrument while measuring could greatly improve the spectral resolution by averaging out sample and magnetic field inhomogeneities [5].

Since the magnetic resonance chemical shift resolution depends on the magnetic field strength, development of instrumentation has progressed towards stronger and stronger magnetic fields, and by the introduction of superconducting magnets in the mid-1970's a leap in field strength was taken compared with the electromagnets previously used. This development increased the physical dimensions of the instruments, and with large magnetic stray fields the instruments required a lot of space and were not suited for work outside the laboratory.

In the late 1980's and the beginning of the 1990's application of pulsed field gradients became routine. Besides enabling a large number of specialised pulse experiments in order to fulfil specific demands by researchers, magnetic resonance imaging (MRI) was also made possible both in the research lab and in hospitals where the technique is normally referred to as "MR scanning". With the introduction of imaging NMR really moved into the field of non-

invasive analysis of large samples, offering totally new prospects for researchers in many fields, and it has become an indispensable diagnostic tool in medicine.

By that time a modern superconducting NMR instrument had become very expensive and expert spectroscopists were required to operate the instruments. Approximately at the same time it was realised that in many research fields the high spectral resolution offered by these instruments was not always required and development of new small bench-top instruments based on the latest electronics and permanent magnets with a much lower magnetic field strength started to emerge. This kind of instruments typically have magnetic field strengths in the range of 0.23 to 0.70 Tesla (T) equal to 10 to 30 MHz for protons. These instruments have particularly found application in the oil industry for rock core analysis, medical diagnostics, food and feed research and bench-top imaging instruments have also found wide application in research. Low-field bench-top instruments have the advantage of being cheaper although still expensive, lighter although still quite heavy and much less sensitive to changes in environment and stray fields whereby a step has been taken in direction of moving NMR into the process line for on- or at-line measurements.

In 1996 a new low-field magnet layout was suggested in a small hand held device called the NMR-MOUSE (MObile Universal Surface Explorer) [6]. Currently, a lot of research is being conducted towards the development of the MOUSE, and NMR has really taken a step towards process application, although this design poses some problems that need to be solved first.

Over the years a number of names have been applied to describe the bench-top NMR instruments such as “low-resolution”, “time-domain”, “wide-line” or “low-field”. Of these only the latter is clearly unambiguous, since a lot of work has been put into producing better magnets and better shimming, thereby producing bench-top instruments capable of showing chemical shift resolution (i.e. “high-resolution” and narrow line width). Furthermore, since high-resolution data may be presented and analysed in the “time domain”, this name is considered equally inappropriate. In this thesis, therefore, “low-field NMR”, or LF-NMR, will be the term used to describe this kind of instrument.

A great deal of development is still being carried out on the NMR hardware, continuously offering new possibilities which make both low-field and high-field NMR indispensable tools in research and routine analysis.

2.2 THEORY

Many nuclei possess a spin angular momentum, which is dependent on the size of the angular momentum quantum number (*I*), which again is dependent on the nucleus in question. The angular momentum quantum number is commonly referred to simply as *spin*.

The spin of a nucleus depends on the mass of the isotope, and nuclei with even mass and even charge numbers possess no spin angular momentum, i.e. $I = 0$. Such nuclei cannot be used for NMR, since it is the nuclear spin property that enables NMR.

When a nucleus that possesses a spin different from zero is placed in a magnetic field, the nucleus will occupy one of a number of energy levels where the number of levels available depends on the value of *I*. The proton (^{1}H) is the most abundant NMR nuclei and has a spin of $I = ½$. For such nuclei there are two different energy levels that the spins can occupy when placed in a magnetic field: $I = -½$ and $I = ½$. This corresponds to an orientation parallel ($I = -½$) or anti-parallel ($I = ½$) with the applied magnetic field. Figure 1 show the two possible energy levels for protons and the dependence of the applied magnetic field B_0. In the rest of this presentation of the NMR theory protons will be used as a model.

Mathematically, the difference in the energy levels shown in Figure 1 can be described as

$$\Delta E_0 = h \cdot n_0 \qquad \text{Eq. 1}$$

where *h* is Plank's constant and n_0 is the frequency of the excitation pulse at which transition between the two energy levels is induced. n_0 is referred to as the *resonance frequency* or the Larmor frequency and will depend on the type of nucleus and magnetic field strength, as described in Eq. 2:

$$n_0 = \frac{g \cdot B_0}{2 \cdot p} \qquad \text{Eq. 2}$$

Here *g* is the gyromagnetic ratio, which is a constant for a given nucleus.

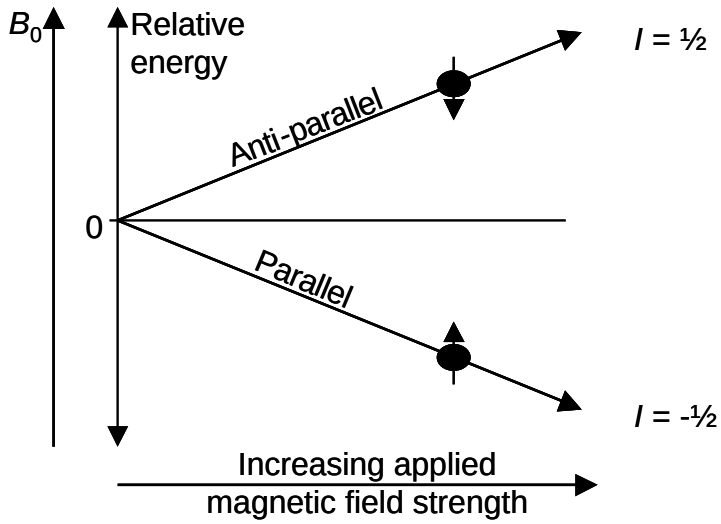

Figure 1 Representation of the two possible energy levels for nuclei with spin = ½ and the dependence of the strength of the applied magnetic field (B_0)

At equilibrium the protons are distributed between the two energy states according to Eq. 3, which is known as the Boltzmann distribution.

$$\frac{N_a}{N_b} = e^{\left(\frac{\Delta E}{k \cdot T}\right)} \qquad \text{Eq. 3}$$

In this equation *a* and *b* represent the spin-up and spin-down states respectively and *k* is the Boltzman constant. According to this distribution there will be an excess of protons in the spin-up state, since this is the energetically more favourable state, and a net magnetisation vector is formed on the basis of this excess. It is this small net magnetisation vector that is utilised in NMR and simple calculations show that at room temperature and a magnetic field strength of 0.47 T (equal to n_0 = 20 MHz for protons) the excess in the spin-up state will be approximately *three protons out of one million*. It is therefore obvious why NMR is commonly described as insensitive compared with other spectroscopic methods, since a large number of protons are required in order to generate an appreciable signal. The low energy levels required for measuring NMR also makes it a non-invasive, low-perturbation method.

While the sample is at equilibrium and unaffected by an external magnetic field, the orientation of the protons will be randomly distributed in all directions. However, when the sample is placed in a magnetic field, the protons will align as described in Eq. 3. Under the influence of the external

magnetic field the protons will start to precess about the direction of the magnetic field, as shown in Figure 2A for a single proton, and the net magnetisation vector pictured in Figure 2B is formed.

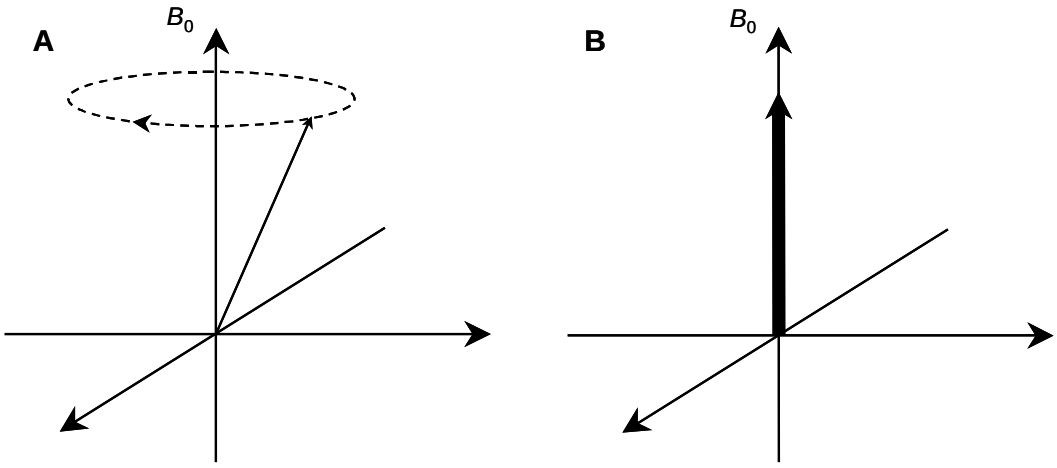

Figure 2 (A) Precession of a single proton around the external magnetic field and (B) resulting net magnetisation vector based on the surplus of protons in the spin-up state.

Since the protons precess at different frequencies due to different local magnetic environment (shielding), they will precess at different rates and fan out, resulting in a net magnetisation vector positioned exactly on the axis parallel to the magnetic field. The bold arrow in Figure 2B represents this net magnetisation vector and is positioned on the direction of the magnetic field with no components on the other axes of the coordinate system.

In order to facilitate description of the spin manipulation that gives rise to the NMR signal it is normal to ascribe axes to the NMR instrument. Thus the z-axis is normally ascribed to the direction of the magnetic field (B_0), the x-axis is the axis along which the excitation radio frequency (RF) pulse is applied and the y-axis is the axis where the signal is detected. Figure 3 shows a diagram of this representation along with the result following a perturbation of the equilibrium system by a RF-pulse.

Strictly speaking, the previous description is not completely correct, as pulses can be given along both the x- and y-axes. Furthermore, the signal is detected along both the x- and y-axes and calculated as the resulting component of these two contributions.

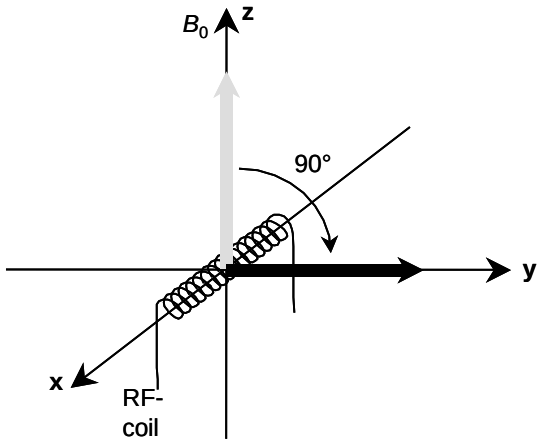

Figure 3 Diagram of the coordinates normally ascribed to the NMR instrument including the RF-coil. The drawing also depicts a perturbation of the system at equilibrium.

In Figure 3 only the net magnetisation vector parallel to the magnetic field is displayed, since as previously described it is only the small surplus of protons that gives rise to the NMR signal. Since the NMR signal is detected as the components along the x- and y-axes, it is clear that a pulse of a given power with a duration precisely long enough to flip the net magnetisation vector into the xy-plane will generate the largest possible signal. This pulse is referred to as a 90-degree pulse and a pulse with a duration long enough to flip the net magnetisation vector along the negative z-axis is referred to as a 180-degree pulse.

After a perturbation by a 90-degree RF-pulse the spin system will lose coherence in the xy-plane, a process known as spin-spin or transverse relaxation, and the loss of coherence is described by a time constant called T_2. This process is due to energy exchange between protons as well as inhomogeneities in the magnetic field, which will particularly influence molecules with mobile protons and high diffusion rates such as water. Simultaneously with the loss of coherence in the xy-plane the protons will seek to regain equilibrium orientation along the z-axis due to the influence of the magnetic field. The time it takes for the protons to regain equilibrium distribution between the two energy states depends on the probability of energy exchanges occurring between the spins and their environment (the lattice). This is characterised by a relaxation mechanism normally referred to as longitudinal or spin-lattice relaxation and is described by a time constant T_1. The nature of T_1 and T_2 relaxation will be described in more detail in the following, along with the pulse experiments commonly used to estimate these time constants.

2.3 PULSE EXPERIMENTS

2.3.1 Free Induction Decay (FID)

One of the simplest pulse experiments possible is an experiment in which a 90-degree RF-pulse is applied and the decreasing signal due to relaxation following the pulse is measured. The loss of coherence and the resulting decaying signal is almost exclusively due to spin-spin relaxation. The FID experiment is schematically described in Figure 4. Note that the timing diagram has not been produced to scale. The duration of a 90-degree pulse is typically in the order of a few microseconds, a 180-degree pulse about twice as long, and the data acquisition time may be as long as seconds. These points also cover the pulse experiments described in the following text.

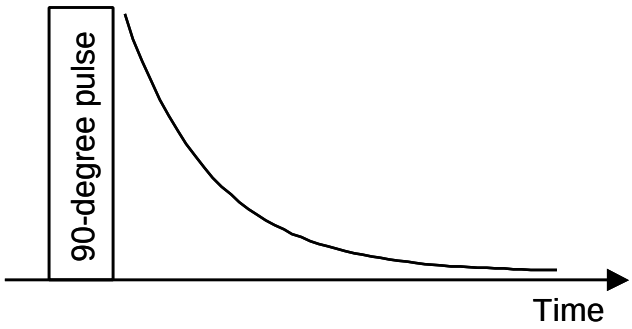

Figure 4 Diagram of the Free Induction Decay pulse experiment

The decay of the signal following the 90-degree pulse depends not only on the relaxation properties of the sample, but also on the homogeneity of the magnetic field. The relaxation time constant measured by a FID experiment is therefore normally shorter than the real spin-spin relaxation time constant and is commonly referred to as T_2^*. For most samples T_2^* is assumed to be much smaller than the true T_2 of the sample.

Following the 90-degree pulse a short delay is required before data acquisition starts due to hardware limitations. However, in the simple FID experiment this delay is shorter than for most other pulse experiments and FID is commonly used for measuring T_2 in samples with a very rapid decay such as most solids. For liquid samples the FID experiment is not commonly used due to the effect imposed on the relaxation decay by magnetic field inhomogeneities and diffusion.

2.3.2 Hahn spin echo

In 1950 Hahn [7] discovered that if a short delay and a subsequent 180-degree pulse followed an initial 90-degree pulse, a spin echo was formed. The reason for the build-up in signal after it has apparently been lost is due to the fact that some of the coherence lost due to spin diffusion is refocused following the 180-degree pulse and the signal is reclaimed.

The delay between the 90- and 180-degree pulses is commonly referred to as tau (τ) and Figure 5 shows a diagram of the Hahn spin echo pulse experiment.

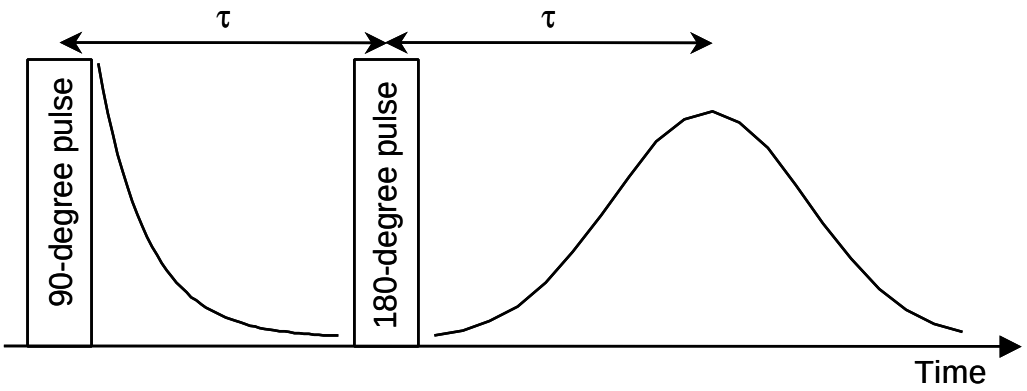

Figure 5 Diagram of the Hahn spin echo experiment

The amplitude of the spin echo will depend not only on the nature of the sample, but also on the length of tau. By making a series of measurements with different values of tau and plotting the spin echo amplitude as a function of tau an exponentially decaying curve will be described. From this curve the T_2 time constant may be found by fitting an exponential to these amplitudes. This T_2 is the true spin-spin relaxation time constant of the sample.

It is important to realise that the 180-degree pulse cannot refocus coherence lost due to magnet inhomogeneities and that the decrease in echo amplitudes is therefore mainly due to the influence of magnet inhomogeneities.

2.3.3 Carr-Purcell-Meiboom-Gill (CPMG)

In 1954 Carr and Purcell [8] suggested an extension of the Hahn spin echo experiment where the T_2 time constant could be determined using only one tau value and the time of acquisition greatly reduced. Meiboom and Gill [9] modified this experiment in 1958, resulting in the pulse experiment commonly referred to as the CPMG pulse experiment.

The CPMG experiment is often preferred to the Hahn spin echo method for measuring the T_2 relaxation time constant since the entire relaxation curve is acquired in "one shot". Just as the Hahn spin echo experiment the CPMG experiment to a great extent compensates for diffusion effects. But more important, by using every second echo only (even echoes), inaccuracies in the 180-degree pulse setting is also corrected for.

. Figure 6 shows a diagram of the CPMG experiment where the first three echoes are shown. Typically, only the top-point of each echo is acquired and the T_2 value of the sample can be obtained by fitting a sum of exponentials to this decaying curve.

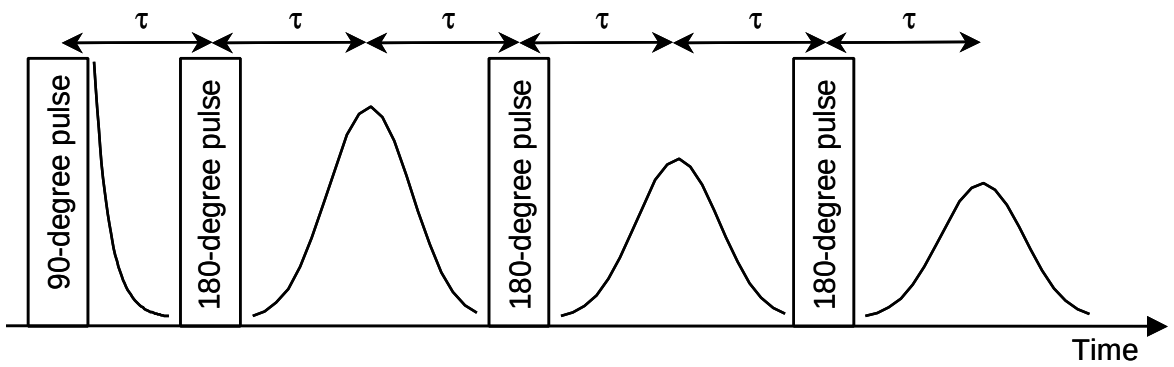

Figure 6 Diagram of the CPMG pulse experiment. Only the first three echoes are shown

There are generally two different ways of depicting and explaining spin manipulation and relaxation. One is the classical approach in which fixed axes are ascribed to the NMR instrument (as seen in Figure 3) and the motions of the spins related to these axes. In the second approach the motions of the spins are related to a coordinate system rotating at the resonance frequency, which is commonly referred to as a *rotating frame*. In Figure 7 the evolution of the spin system in the CPMG experiment is described in more detail in a rotating frame showing only the effect of the first couple of 180-degree refocusing pulses.

It is seen that the net magnetisation vector is split into two parts (dotted arrows and dashed arrows) rotating away from or against each other, depending on the location in the pulse experiment. These two parts represent spins moving more slowly (dots) or faster (dashed) than the resonance frequency due to local differences in the magnetic field strength of the lattice.

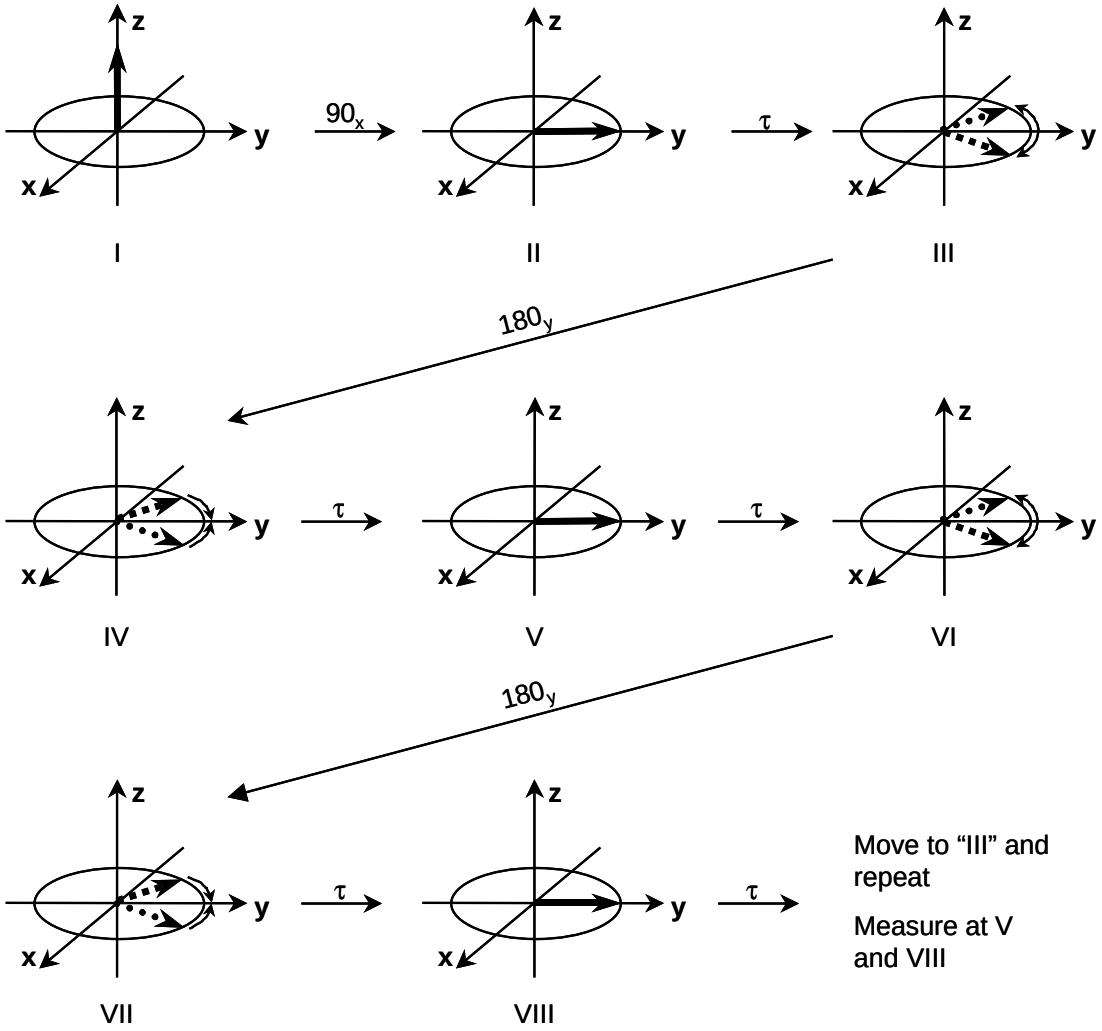

Figure 7 Diagram of the evolution of the spin ensample in the CPMG pulse experiment. The figure shows de-phasing and re-phasing following the 180-degree refocusing pulse.

2.3.4 Inversion Recovery (INVREC)

The inversion recovery pulse experiment [10] is typically performed in order to estimate the spin-lattice relaxation time constant (T_1) which describes the time it takes for the protons to return to equilibrium distribution along the magnetic field following a perturbation. Due to the fact that signal can only be detected in the xy-plane and not along the z-axis, spin manipulation is required in order to facilitate measurement of T_1.

Although quite simple to perform the inversion recovery pulse experiment is somewhat more difficult to describe than the previous experiments. The experiment consists of an initial 180-degree pulse and a recovery delay (D1) followed by a 90-degree pulse in order to flip the net magnetisation vector from along the z-axis and into the xy-plane to measure the current size of the

net magnetisation vector. The experiment is repeated for a range of recovery delays and for each D1-value a point is generated on the final relaxation curve plotted as a function of D1. The resulting relaxation curve can be exponentially fitted and the relaxation time constant estimated. It can be seen that the procedure of data acquisition closely resembles the measurement of T_2 by the Hahn spin echo experiment previously described.

The experiment is quite time-consuming and is therefore mainly used for determination of T_1 relaxation time constants and not so much for predictive purposes, although excellent predictive performance has been achieved using T_1 data [P1,11,12]. In Figure 8 the evolution of the spin system is presented in the rotating frame for five different recovery delays (D1) prior to the 90-degree pulse.

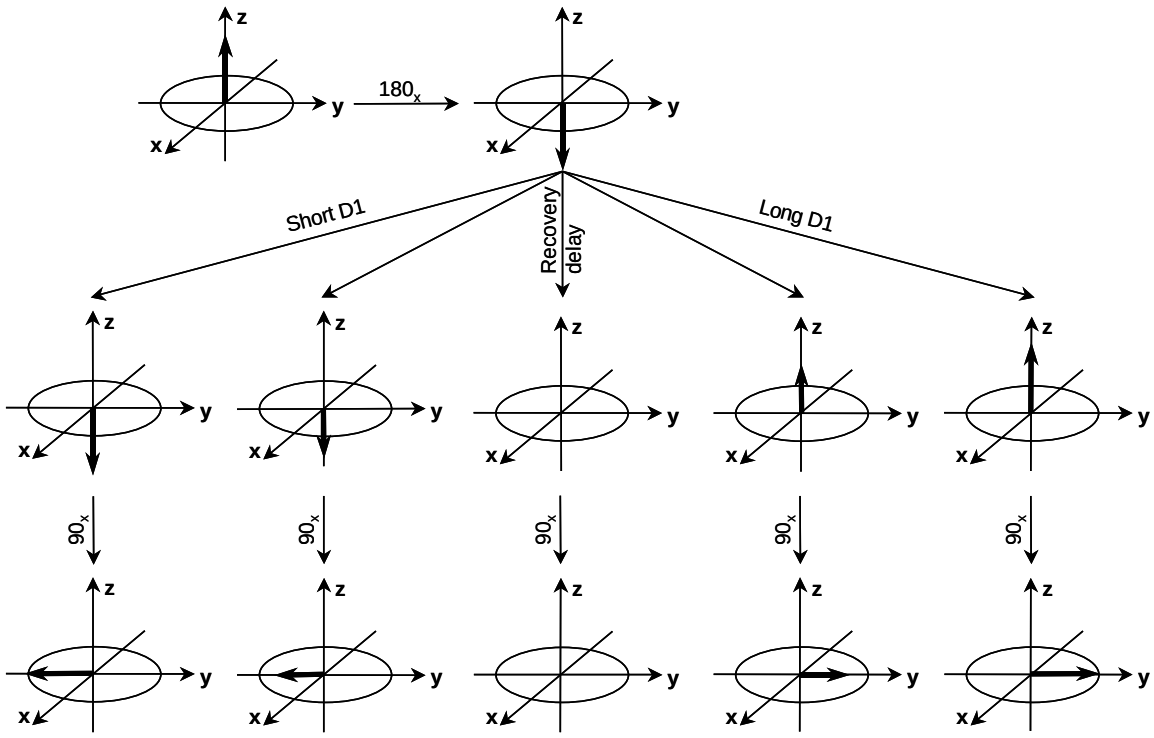

Figure 8 Diagram of the evolution of the spin ensample in the rotating frame for the inversion recovery pulse experiment. The figure shows the situation for five different values of D1.

The five situations correspond to increasing D1-values starting with a very short value to the left, a long D1-value to the right and in the middle the situation in which D1 is chosen to be equal to the value where the net magnetisation vector is equally distributed between the spin-up and spin-down states (D1 = $\ln(2)\cdot T_1$) and the signal therefore completely disappears. The second and fourth situations are simply D1-values in-between.

3 LF-NMR AND DATA ANALYSIS

3.1 TRADITIONAL METHODS AND APPLICATIONS

A range of pulse experiments in addition to those just described in Chapter 2 may be applied. There are furthermore a number of different ways to analyse the acquired LF-NMR data. Some of these possibilities will be described in the following.

The importance of LF-NMR in food science arises from the fact that food samples with complex structures can readily be analysed. Food samples are typically composed of many different components such as carbohydrates, proteins, water, fats and other minor components, and food samples are normally considered to be very inhomogeneous. Besides measuring chemical composition of a sample, the physical properties such as the states of water may be investigated, completing the range of information obtained to produce a clearer and better understanding of the samples analysed. This is of outmost importance since the states of water controls a great number of properties in food materials, such as rheology, sensory properties, stability to oxidation, microbiological activity and many more.

The problem of sample inhomogeneity is readily handled by LF-NMR, measuring on a volume of sample rather than just the surface of the samples, and as an additional advantage LF-NMR is normally considered to be non-invasive and non-destructive.

Signal generated by the FID pulse experiment originates from all protons in the sample and the initial signal intensity of the relaxation decay is therefore proportional to the total number of protons in the sample. However, due to hardware limitations there is normally a delay following the 90-degree pulse before acquisition of data can commence. Because of this delay (commonly called “dead time”), it is not possible to acquire the true signal originating from all protons, as will already have relaxed at the time when signal acquisition starts.

The relaxation pattern of the protons in a sample will depend on the nature of the protons, e.g. protons in water or in lipids, as well as the state of the protons, e.g. structured or free water is important. Solid materials typically have a very short T_2, whereas liquid materials have longer T_2 time constants.

It is these differences that are utilised when setting up an experiment as well as in the subsequent data analysis.

One application that directly takes advantage of the different relaxation patterns of solid and liquid protons and which has been widely used in the industry is the determination of solid fat content (SFC) in fats and oils [13,14,15]. Determination of SFC has traditionally been achieved by assuming that the amplitude of the first measured point corresponds to the entire proton content in the sample (solid + liquid) and that a point acquired some time later is assumed to correspond to liquid component only. The assumption is that since the solid protons decay rapidly, waiting a certain length of time will leave only the liquid signal to be acquired. The first point is typically collected 11 μs after the 90-degree pulse (due to the "dead time") and the second point has typically been acquired approximately 70 μs after the pulse, at which time it is assumed that only the liquid oil signal is left. Based on these two points it is now simple to calculate the solid fat content of the sample. A similar approach has been used in a number of different applications such as the determination of the oil content in oil seeds [16,17].

By using the Hahn spin echo pulse experiment and acquiring the initial point in the FID as well as the top point of the spin echo following the 180-degree refocusing pulse a tremendous improvement of the method just described can been achieved. Still, the initial point in the FID is related to the total proton content, and since the fast relaxing components will not be refocused if tau is chosen appropriately, the spin echo amplitude will correspond to the remaining slow relaxing components. With this set-up SFC could be determined with better precision. However, one of the main achievements of this approach was that now the fat content could be determined in meat samples where a small amount of water was present. It thus became possible to measure the fat content without drying the sample prior to the NMR measurement [18,19,20,21].

As already mentioned, data originating from CPMG, INVREC or even the Hahn spin echo experiment repeated for different tau values produce decay curves that can be described by sums of exponentials. This relationship can be seen in Eq. 4 and Eq. 5 for T_1- and T_2-relaxation respectively and has been widely used for qualitative as well as quantitative purposes.

$$I(t) = \sum_{i=1}^{N} M_{0,i} \cdot \left(1 - 2 \cdot \exp\left(\frac{-t}{T_{1,i}}\right)\right) \quad \text{Eq. 4}$$

$$I(t) = \sum_{i=1}^{N} M_{0,i} \cdot \exp\left(\frac{-t}{T_{2,i}}\right) \quad \text{Eq. 5}$$

Here *I(t)* describes the signal intensity as a sum of *N* mono-exponentials with $M_{0,i}$ and $T_{x,i}$ expressing the initial signal amplitude and relaxation time constants respectively. By using a computer and an appropriate algorithm it is possible to calculate amplitudes and time constants for a given number of components and use these parameters to characterise the sample. The amplitudes of the resulting fit mainly provide quantitative information, as they describe the amount of the different components, while the corresponding time constants mainly give qualitative information, since they relate to the origin or state of the protons of the given component. To use this analysis properly, knowledge is required as to how many components can be expected in the sample and time constants for different pure components need to be known in order to infer the nature of the deconvoluted components.

With the introduction of the computer it became possible to apply exponential fitting routinely and the procedure represents a cornerstone of traditional data analysis of LF-NMR data. Based on the amplitudes it now became possible to simultaneously determine oil and water content in a variety of more complex samples.

Exponential fitting has been applied in a wide range of applications such as monitoring of alcoholic fermentation [22], different applications on grain [23,24], determination of sublimation endpoint in freeze-drying [25], examination of textural changes in frozen cod as well as stored or processed cod [26,27], measurement of mobility of lipids in bread [28], determination of glass transition temperature in food polymers [29], and determination of iodine number [30,31].

In the late 1970’s discussions started as to whether it was correct to interpret relaxation curves acquired on complex samples by simple exponential fitting, describing the sample by two or three underlying components. Clearly, exponential fitting gives useful information and it is an easy way of analysing samples, but even in a sample with only two components interactions at

boundaries between the two components, exchange effect and other effects will result not only in two time constants but in a distribution of time constants and thus in a large number of components. It was therefore argued that relaxation curves should be fitted by a distribution of exponentials, i.e. by a much larger number of "components" than the normal few. Some of the first work presenting distributed exponential fitting as an analytical tool was carried out by Provencher and Dovi [32] in which photodissociation of CO from heme proteins at low temperatures in simulated data is described, Lillford *et al.* [33] used distributed exponential fitting of LF-NMR data, to describe the distribution of water in meat, showing that improved interpretation results when distributed exponential fitting is used rather than discrete exponential fitting.

Although not used in as many applications as discrete exponential fitting, the distributed exponential fitting has still been applied for a number of different purposes such as the study of changes in sub-cellular water compartmentalisation in apple tissue during drying and freezing [34], quantitative determination of bound water in wheat starch [35], and the study of the state of water and oil in frozen emulsion [36]. In a recent paper by Bertram *et al.* [37] the analysis is also used to describe the distribution of water in meat as a function of different treatments in a comparison of two traditional methods for determining WHC as well as estimated by LF-NMR.

By comparing the results of discrete and distributed exponential fitting performed on the same samples it is found that time constants calculated in the discrete analysis often coincide with the peaks resulting from distributed exponential analysis. This clearly shows that the distributed exponential analysis, although a simplification, does in fact describe information relevant to the underlying features of the samples.

Mathematically speaking, the problem of fitting a distribution of mono-exponentials to a relaxation decay is said to be ill-posed, meaning that there exist no simple solution to the problem and additional knowledge and constraints are required to find a solution. The mathematical term for transforming an ill-posed problem into a problem that can be solved is called *regularisation*, which is a research field of its own [38]. A number of

publications have been made using existing algorithms or suggesting new algorithms to improve the distributed exponential analysis [39,40,41,42].

It is important to realise that although distributed exponential fitting offers a nice way for interpreting LF-NMR data in relation to changes in the distribution of water or fat or some other constituent, some problems need to be taken into account before applying the method. These relate in particular to the fact that different sample characteristics and noise levels may severely distort the obtained results. Caution should therefore be taken when distributed exponential analysis is applied to a set of different samples with the purpose of comparison. Application to a data set based on repeated acquisitions from a single sample, where the different acquisitions result from, for example a change over time, seems to be more appropriate. Distributed exponential fitting will be further described in a later section.

In 1992 Davenel and Marchal [43] described an improvement of a study reported by Tollner and Hung in 1990 [44]. Davenel and Marchal described a problem from the feed industry where a number of different raw materials with different moisture content are mixed to produce a feed product.

In this paper Davenel and Marchal applied canonical factor analysis to the entire data set in an attempt to classify the different raw materials. This is probably one of the first applications of multivariate data analysis within LF-NMR on food materials. In the rest of this chapter and the next a more detailed description of some of the most common chemometric methods applied and the results obtained will be discussed.

3.2 EXPLORATIVE METHODS

3.2.1 Univariate

Analysis of variance (ANOVA) is a tool used to investigate whether or not acquired data contain information relevant to expected groupings in data. In a data set containing one relaxation profile for each sample the analysis is applied to the individual variables one after the other and in the end the calculated variances are plotted as a function of the position in the signal. This “variation-curve” will contain the same number of elements as the acquired data and by comparing the two plots the variables containing the most information regarding the problem in question can be selected. If additional

univariate tools are applied in the further analysis the best variables for the specific problem can be selected on the basis of the result of the ANOVA.

The analysis consists of calculating two different variance measures, one describing the difference between the groups examined (V_G) and one describing the residual variance (V_R). The equations can be written as:

$$V_G = \frac{\sum_{j=1}^{g} n_j (\bar{x}_j - \bar{\bar{x}})^2}{(g-1)} \qquad \text{Eq. 6}$$

$$V_R = \frac{\sum_{j=1}^{g} \sum_{i=1}^{n_j} (x_{ji} - \bar{x}_j)^2}{\sum_{j=1}^{g} (n_j - 1)} \qquad \text{Eq. 7}$$

where g is the number of groups, n_j is the number of samples in group j, x_{ji} is the value for sample i in group j, $\bar{x}_j$ is the mean value for group j and $\bar{\bar{x}}$ is the grand mean value.

ANOVA has been applied in a series of publications using LF-NMR for the analysis of different problems such as characterisation of cocoa masses [45], for examination of spreads and gelatines with different composition, for mixtures of cation (Cu^{2+}) and a ligand (tetraphenylporphin), and for glycine solutions at different pH values [46,47,48] as well as for technological meat quality parameters [49].

In the mentioned studies ANOVA is applied to INVREC data, to CPMG data and to data comprised of the two, either by acquiring true T_1-weighted CPMG data or by calculation of the outer product (see section 3.2.2 below). The data acquired by T_1-weighted CPMG [47,48] or calculated by the outer product in reality describe a two-dimensional (2D) landscape for each sample which has to be unfolded (vectorised) before ANOVA is performed. Once information in data relevant to the problem in question has been confirmed by the ANOVA further predictive analysis is performed in which the entire relaxation profiles can possibly be used in a multivariate way.

A drawback of ANOVA is that due to the univariate nature of the method interactions between the different variables are not taken into account, even

though it is known that LF-NMR data are highly correlated and that interactions between variables can be expected.

3.2.2 Multivariate

In spectroscopic signals it is normally expected that the variables within the spectrum are strongly correlated, and LF-NMR data is no exception, as demonstrated by Bechmann *et al.* [P2]. Here it is shown that more than 99 % of the variables have an intercorrelation larger than 0.9. Due to these high intercorrelations multivariate techniques must be preferred for analysis of this kind of data, since these methods can handle large data sets with correlated variables, and interaction effects are taken into account in the resulting models.

One multivariate method is Principal Component Analysis (PCA) [50,51] which is a decomposition method where a set of corresponding scores (**T**) and loadings (**P**) is calculated for a data set (**X**), leaving only a residual (**E**):

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P}^{\mathrm{T}} + \mathbf{E} \qquad \text{Eq. 8}$$

Scores correspond to concentrations of the common latent variables (loadings) as displayed in Figure 9, where the decomposition of three samples chosen at random from a PCA analysis of a LF-NMR data set is shown.

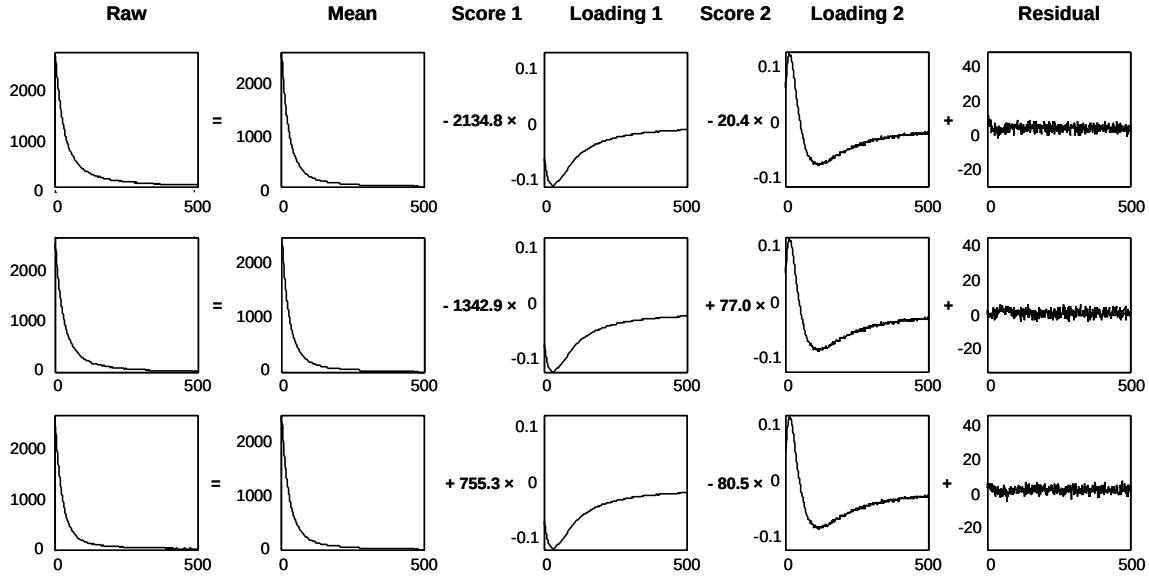

Figure 9 Plot of three samples decomposed by PCA into a common mean spectrum and two sets of corresponding scores and loadings as well as a residual.

The idea of PCA is to find orthogonal variations present in the data set on a model-free basis until only noise is left in the residual. The variations are extracted in decreasing order with the largest variation first called the first principal component and so forth. PCA is a very strong tool used to explore the variation present in the data set and often the origin of the extracted variances can be interpreted directly through plots of the calculated scores by utilising known information about the samples.

Another valuable tool offered by PCA is outlier detection. Outliers are samples that do not conform to the majority of samples in a data set and will therefore have a detrimental effect on the calculated model. The reason for a sample being an outlier can be one of many, such as wrong sample preparation or erroneous handling of the sample, but it can also be due to instrumental artefacts. It is important to make an effort to find out why a sample is an outlier before removing it from the data set. Once a sample has been shown to be an outlier, it should be removed from the data set and the model must be recalculated.

One of the first applications of PCA to LF-NMR on food materials is by Hernandez and Rutledge [45] where the influence of supplier and roasting of cocoa beans is evaluated. Another good example of the explorative ability of PCA is shown in Jepsen *et al.* [P1] where it is seen that the first principal component describes water content in a data set consisting of measurements of cod and salmon and it is also seen that cod and salmon samples can clearly be separated. In Pedersen *et al.* [P3] two different varieties of seeds as well as different water content within each variety of seeds can be visualised. Thybo *et al.* [52] applied PCA to evaluate the relationship between sensory evaluation of different potato samples treated differently and Engelsen *et al.* [P4] visualise the changes in proton mobility taking place during baking of a small bread directly inside the NMR probe with data continually being acquired during the time of baking.

Another approach for enhancing the information content of spectral data was suggested by Rutledge *et al.* [46,48,53] based on the calculation of the outer product of two acquired signals of different nature. The idea is not only to enhance interpretation through ANOVA, but also to enhance the predictive ability by combining and thus simultaneously treating the two different

signals. In the papers it is concluded that enhanced interpretation results from the procedure when compared to analysis of the data sets for the individual signals, and in one of the studies it is concluded that although it is clear that outer product data generated with INVREC and CPMG data do not contain the same information as the true T_1-weighted CPMG data, the data structure generated by forming the outer product represents a good approximation [48].

The outer product is calculated and understood as displayed in Figure 10 in an example in which constructed INVREC and CPMG data are used. The subplots (A) and (B) show the pure relaxation profiles for INVREC and CPMG respectively, and (C) shows the calculated 2D landscape resulting from calculating the outer product of the pure signals. In (D) this landscape is unfolded to produce a vector that is used in the following analysis.

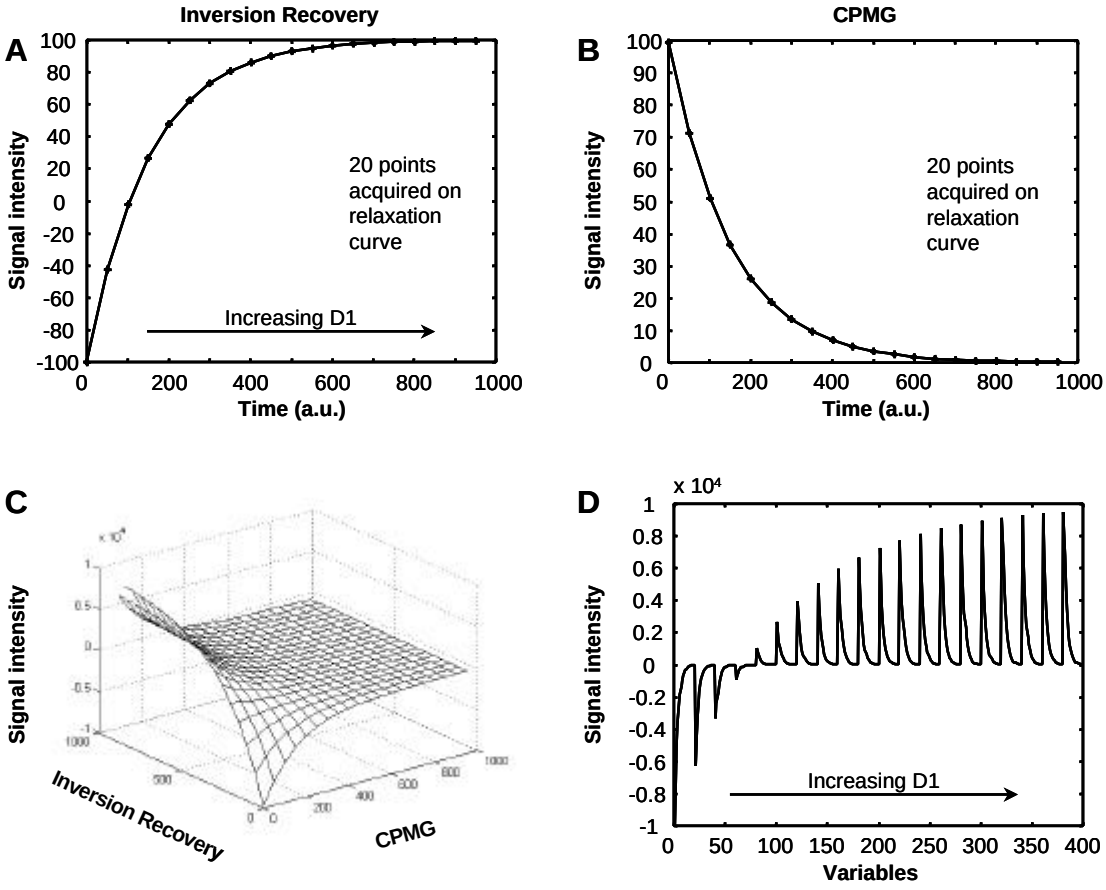

Figure 10 Illustration of the outer product (C) calculated from INVREC data (A) and CPMG data (B) as well as subsequent unfolding for further data analysis (D)

At first sight the idea is quite interesting and the conclusions presented in the papers convincing. From a mathematical point of view there is, however, a problem with the approach, which is described in the following:

> Assume that a sample **X** contains a true 2D T_1-weighted CPMG data matrix originating from one sample. If INVREC and CPMG data both are sums of F underlying exponential components, then **X** can be described as:
>
> $$\mathbf{X} = \mathbf{S} \cdot \mathbf{C} \cdot \mathbf{N}^T \qquad \text{Eq. 9}$$
>
> where **S** represents the pure underlying profiles in INVREC data and has dimension I×F (I data points acquired), **N** represents the pure underlying profiles in CPMG data with dimension J×F (J data points acquired) and **C** is an F×F diagonal matrix of concentrations. The matrices **S**, **C** and **N** all have rank equal to F, and **X** will therefore have rank F meaning that F independent phenomena are reflected in the measured data.
>
> Now assume that a pure CPMG profile has been acquired for the same sample without T_1-weighting, the acquired signal can be represented by **n** with dimension I×1 and similarly for INVREC the pure profile is given by **s** with dimension J×1. The outer product of **s** and **n** is equal to $\mathbf{s}\cdot\mathbf{n}^T$, which will have rank one.
>
> Due to the fact that **X** have rank F, proving that **X** is different from $\mathbf{s}\cdot\mathbf{n}^T$, it is possible to derive the information in $\mathbf{s}\cdot\mathbf{n}^T$ from **X** (by simply taking the appropriate row and column and multiply these). However, it is not possible to obtain the information in **X** from $\mathbf{s}\cdot\mathbf{n}^T$ and hence more information is obtained from **X** than from $\mathbf{s}\cdot\mathbf{n}^T$. Thus, the outer product data cannot increase the information content and replace true 2D data.
>
> Enhanced interpretation by forming the outer product is also unlikely, as the outer product is simply a linear combination of the original profiles and no physical interaction between the two signals will thus be expressed.

On this basis it is concluded that outer product cannot enhance the information content either through ANOVA analysis or through PLS

prediction models. However, the true 2D data, for example acquired by a T_1-weighted CPMG pulse experiment may contain more information than the two separated data sets together, since in this case the influence of the T_1-weighting on the individual components in the CPMG relaxation signal may be different and interaction effects will thus be present.

It should be stressed, that in the example given above it is assumed that both INVREC and CPMG data consist of the same number of components, but that this is unimportant, since the basic content of the argument is equally valid if the two LF-NMR data contain different numbers of components.

3.3 CALIBRATION

3.3.1 Univariate

Typically, a single response variable is measured or a ratio calculated between two measured or calculated responses. Based on this single variable (ratio) for each sample linear regression is performed and from the calculated regression line a correlation, a prediction error and a bias (a measure of the average difference between measured and predicted values) can be calculated, giving indication of how well the measured variables describe the reference variable in question.

There are a few problems that should be considered when using univariate predictions based on linear regression. One is that out of a large number of measured variables only one variable is selected for the linear regression. It is thus crucial to select the measured variable that best correlates with the reference variable, and it is here that tools such as ANOVA or another kind of variable selection comes into question. Besides ANOVA, different schemes for variable selection may be applied, such as forward selection [P2].

Another problem occurs if different components in the sample give signals in the same region or a variable baseline offset is possible in data. In cases of coinciding signals these will be inseparable when only a single variable is measured and the content will be estimated too high due to the several components adding to the signal intensity. Regarding the changing baseline offset an example is given in Figure 11 that clearly explains the nature of this problem. The data described in Figure 11 are not directly related to a problem seen in LF-NMR, since normally no offset is present. What might be seen in

LF-NMR is, for example, that since in the spin echo resulting from a Hahn spin echo experiment most protons contribute to the acquired signal, the desired component may be masked. Or in CPMG relaxation data the problem becomes apparent if the time constants of the different components are too similar to be resolved by the applied algorithm. In this case it will not be possible for the exponential fitting algorithms to separate the components.

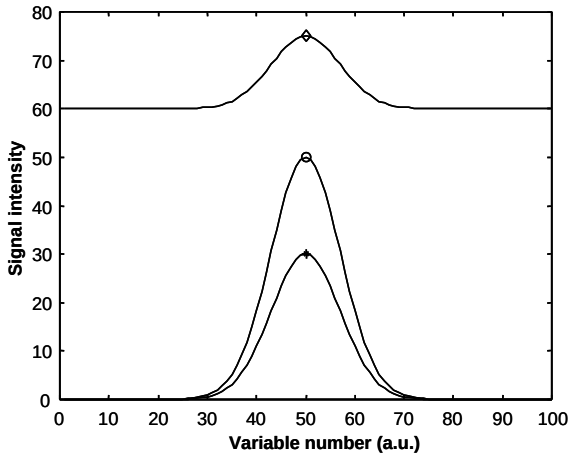

Figure 11 Example of the problem with variable offset if univariate linear regression is applied

On the plot it is obvious that if only the centre point of the peaks is acquired the curve marked with "◊" will clearly be estimated to have the highest content of whatever compound is in question, even though this is clearly the response with the smallest amplitude relative to the signal baseline. For this reason univariate linear regression based on interacting or coinciding signals cannot be expected to produce useful and trustworthy results.

3.3.2 Multivariate

Measurement of more than one response variable may be used for multivariate predictive purposes, in a similar was as it has previously been shown for PCA. A number of different possibilities exist for calculation of multivariate prediction models where the simplest approach is a multivariate extension of the univariate linear regression, which has been given the obvious name of Multiple Linear Regression (MLR). MLR calculates the direct correlation to the reference variable and it is required that the number of variables used is smaller than the number of samples present in the data set. Rather than using a few measured or calculated variables it is possible to use scores for a given number of components calculated by PCA as the regression

variables in a MLR. This approach is called Principal Component Regression (PCR). The advantage of this approach compared to MLR is that the entire variation in the data set is compressed by PCA into a few variables (scores) and the entire variation in the data set is thus used in the prediction model.

Another possibility is to use Partial Least Squares (PLS) regression [54,55,56] where the variation between the data matrix and the reference variable is maximised through an iterative procedure. In this way the variation in the measured data matrix directly correlating with the variation in the reference variable is extracted. The variation structure calculated by PLS is not necessarily identical to that of PCA. This is due to the fact that in PCA the data matrix is decomposed independently of the reference value, whereas in PLS the data matrix is decomposed to optimally describe the reference value. Just as for PCA, detection of outliers is a very important feature in PLS.

If, for instance, in the example given in Figure 11 the entire response curve is acquired and used in a multivariate prediction model, for example, by PLS, the algorithm should be able to extract the relevant information and eliminate the irrelevant baseline offset information. This means that the amplitude relative to the baseline rather than the absolute signal intensity is modelled and the correct order of the profiles will be obtained.

A large number of successful applications of especially PLS on LF-NMR data have been reported in the literature with the first work published in 1996 by Thygesen [57] in an application where the moisture content and basic density of softwood was determined. The following year, the first application of LF-NMR to food materials was presented by Gerbanowski *et al.* [58] for the determination of moisture in meat products. Over and over again it has been shown that multivariate regression by PLS on the entire relaxation decays is superior to simple univariate prediction models or models based on MLR on, for example, concentrations and/or time constants calculated by exponential fitting [P1,P6].

3.4 CURVE FITTING

In LF-NMR two fundamentally different approaches for curve fitting are applied. The simplest is discrete exponential fitting where data are approximated by a sum of a few mono-exponential components. The other is

distributed exponential fitting where data are approximated by a distribution of mono-exponential components rather than just a few components. Both procedures take advantage of the fact that data acquired by INVREC and CPMG pulse experiments can be described by sums of exponentials, as indicated in Eq. 4 and Eq. 5 respectively on page 19.

Traditionally, exponential fitting has been used for two purposes: one for qualitative and the other for quantitative interpretation. The qualitative character relates to the fact that the time constants to some extent reflect the nature or state of the protons in the sample. Thus, changes in time constants from one sample to another can be used to draw conclusions with respect to changes concerning the state of the protons. The quantitative nature of the exponential fit comes from the fact that the calculated amplitudes reflect the amounts of the different components in the sample, and again changes in amplitudes provide information on the nature of the changes taking place. In addition the amplitudes are often used either to calculate a ratio for prediction by linear regression or all amplitudes (and possibly also the time constants) are used simultaneously for prediction purposes using MLR. A large number of references to such application can be listed, but only a few will be given here [P1,52,58,59,60].

Since the first treatment of exponential curve fitting in 1795 [61] a great number of different algorithms to perform this task have been proposed [40,62,63,64,65]. In a paper by Bechmann *et al.* [P2] an algorithm is described where the fact that the parameters to fit – N sets of amplitudes and time constants – are separable is utilised. This means that estimation of the non-linear time constants and the linear amplitudes can be treated as two individual problems, and in the described algorithm the time constants are found by a Simplex minimisation followed by a least squares fit of the amplitudes inside the function evaluation call. This procedure may not be the fastest possible; however, it has turned out to be very robust.

One of the problems with discrete exponential fitting is the fact that the time constants will shift as a function of interaction effects. Moreover, the time constants will also shift depending on the concentrations of the different components. This is problematic, as interpretation will be wrong if the change in time constant is ascribed to a shift in properties when it is only a

consequence of changes in concentrations. One way to overcome this problem is to fit all samples in a data set simultaneously, thus calculating common exponential components. This will produce a decomposition similar to that, which is seen for PCA, where common latent variables are calculated and all the variation between samples is moved into the concentrations. In this approach, however, loadings will be strictly exponential and not orthogonal – both factors making the procedure conceptually different from PCA.

This approach has previously been described by Wijnaendts van Resandt *et al.* [66], and in the paper by Bechmann *et al.* [P2] non-negativity constrained alternating least squares regression (NN-ALSR) on the entire matrix is also applied to solve this problem. Unfortunately, NN-ALSR being iterative is very slow and stability problems occur if too many components are extracted. It has later been shown that the algorithm for discrete exponential fitting proposed in the paper by Bechmann *et al.* [P2] can easily be extended to also calculate the fit to the entire data matrix, giving a robust algorithm for this purpose [P7].

In order for this approach of exponentially fitting all samples in the data set simultaneously to be valid, identical time constants must be assumed for all samples. This means that data sets with samples containing different time constants, such as data resulting from heating of a sample during continuous measurement, will most likely not be valid.

It is expected that this approach will greatly enhance both interpretation and prediction of future data sets by stabilising the obtained fit and by removing drifting time constants due to algorithmic instabilities.

As previously described, most food materials are considered rather complex and inhomogeneous on a macroscopic level and it has been argued that it is a crude approximation of reality to describe these samples by a sum of only a few discrete exponential components. It has therefore been proposed to use distributed exponential fitting for better interpretation. It can, however, be shown mathematically that the problem solved by distributed exponential fitting is ill conditioned and a simple solution cannot be found without applying constraints, i.e. regularising the problem. This means that there are a multitude of solutions given as distribution curves that will all produce the

same or very similar fit to the data, however resulting in totally different interpretations. It is quite simple to visualise this problem mathematically, but this is also an easy task graphically, as shown in Figure 12 for a simple example using constructed data.

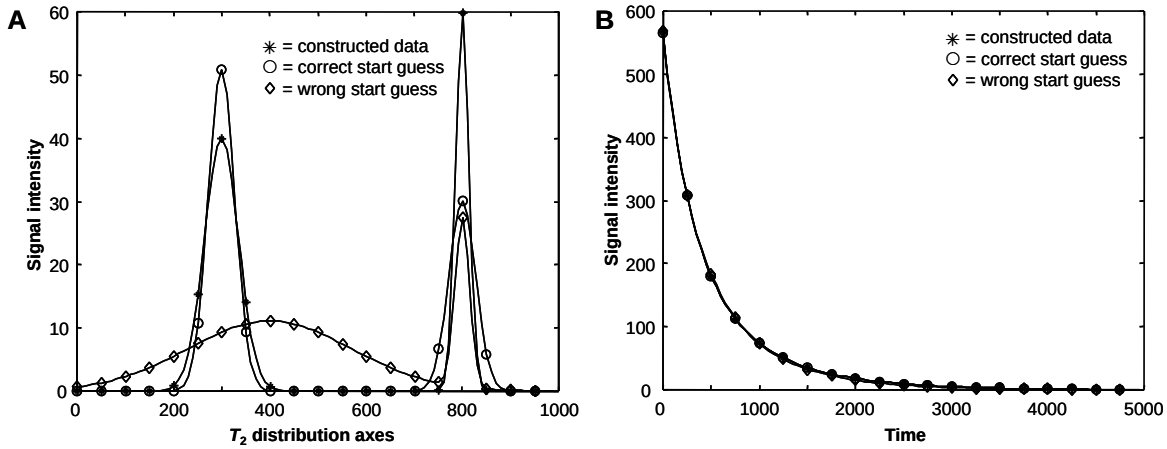

Figure 12 Display of the nature of the ill conditioning of the problem to be solved by distributed exponential fitting. (A) shows the original and two calculated distributions, while (B) shows the corresponding three relaxation profiles. The profile marked with "✳" is the original generated profile, "○" represents an unconstrained fit with the correct peak centre values entered as start guesses and "◊" represents an unconstrained fit with wrong peak centre value as start guesses for one of the two parameters.

Now the relaxation curves corresponding to distributions for "○" and "◊" are calculated (B) and the error of the fit calculated by comparing the obtained fits with the original true relaxation profile. The squared sum of error (SSE) of the fit corresponding to "○" is 0.83 and SSE for the second fit corresponding to "◊" is equal to 18.76.

From the SSE one would think that the two calculated distributions should result in very different relaxation curves with such a difference in residual error. However, when the three relaxation curves are compared they appear to be extremely similar (B). The distributions, on the other hand, are very dissimilar. It is thus obvious that no simple solution can be found.

From this example of an unconstrained algorithm it is obvious that constraints must be applied in order to stabilise the solution, and the problem is to figure out what constraints to apply, since no knowledge really exists concerning the shape of the underlying profiles. Furthermore, it has just been demonstrated that appropriate start guesses for the parameters involved are very important.

Unpublished work using a commercially available program has shown that several parameters, such as noise level as well as whether or not a sufficient number of data points have been acquired in order to include a region of baseline, will significantly influence the calculated distributions.

In an attempt to suggest a new algorithm (unpublished work) it was proposed to use the regression vector from a PLS model for a given number of components. In this model pure T_2-profiles corresponding to a selected distribution range make up the **X** matrix and the acquired relaxation profile of the sample to be fitted make up the **Y** vector. The idea is that PLS maximises the variation in **X** corresponding to the variation in **Y** and therefore the regression vector might be a good start guess for the distribution. Since the regression vector in PLS is not constrained to non-negativity, a subsequent non-negative least squares smoothing was applied to the regression vector, in the hope that this would produce a correct and stable distribution. It turned out that the convergence of the smoothing step would require additional constraints in order to be consistent and also the distribution did not converge to the *a priori* known solution in generated data. For this reason this work has been discontinued for the time being.

Contrary to discrete exponential fitting, distributed exponential fitting has only been used for interpretation purposes. Prediction has been left to univariate techniques and multivariate techniques based on calculated parameters or directly on the acquired relaxation spectra [46,48,49,67].

3.5 VALIDATION OF MODELS

Validation of the calculated models is an important aspect both for exploratory and predictive data analysis. For explorative methods such as PCA, validation is important in order to make sure that the interpretation is not based on a wrong assumption regarding the number of components present in the data set. For multivariate predictive methods the aim of validation is again to make certain that the model is based on the correct number of components so that noise is not included in the prediction, but it is also important to make sure the resulting estimate of the prediction error is as close to the true value as possible.

Based on the number of samples in a data set as well as the nature of the samples present, a range of different validation schemes exists, of which only two good and often used methods will briefly be mentioned: cross validation and test set validation which are both used in predictive modelling. It should be mentioned that there is much controversy as to which method is best, and the current presentation is not intended to be an active part of this discussion.

3.5.1 Cross validation

Cross validation is often used in the case where a limited number of samples exist. In general terms the data matrix is divided into *N* segments with *M* samples in each segment and one by one a segment is excluded, a full prediction model calculated on the remaining *N*-1 segments, the excluded segment is predicted in the calculated model and the error calculated by comparing the predicted values to the known reference values. In the end the error for all the different segments is average and the total error for the given number of components is obtained. Quite often *M* is equal to 1, meaning that there is only one sample in each segment, which in turn means that *N* is equal to the number of samples in the data set. This situation is commonly referred to as *full cross validation*.

The method is somewhat time-consuming, since a full model needs to be calculated for each segment, but the method is good and has been applied in numerous papers where multivariate predictive analysis is applied [P3,P6,12,52].

3.5.2 Test set validation

Test set validation is often used for data sets in which the number of samples is large enough to divide the samples into two subsets – one for calculating the prediction model, and the other for estimating the prediction error. One may refer to a dependent test set, which is taken from the original set of samples, or an independent test set where a new series of samples is analysed and used as test set.

When test set validation is performed it is important to make certain that the two subsets span the same range of variations in order to ensure that the calibration model corresponds to and spans the variation in the prediction set. This way the most correct estimate of the prediction error is obtained.

Quite often, when test set validation is applied for estimation of the prediction error, cross validation is applied initially when the number of components to be used in the model is estimated, and subsequently the prediction error is calculated using the test set. Several papers have been published using test set validation, a few of which are mentioned [P1,11,58].

3.5.3 Durbin-Watson

A different statistical approach for estimating the number of components to be used in a multivariate prediction model has been used and described by Rutledge and co-workers [47,48,53]. In this test the structure or "non-randomness" in the residuals, for instance, after a PLS prediction model, is calculated. The method is called the Durbin-Watson test and is calculated as described in Eq. 10.

$$D = \frac{\sum_{i=2}^{n} (dx_i - dx_{i-1})^2}{\sum_{i=2}^{n} (dx_i)^2} \qquad \text{Eq. 10}$$

where dx_i and dx_{i-1} are the residuals for n successive points in a series. It can be calculated that for $n > 100$ the distribution will be random with a 95% confidence interval when the Durbin-Watson D-value lies in the range from 1.7 to 2.3 [48].

3.5.4 JackKnife

Estimation of the number of components to fit when applying multi-exponential fitting also poses a problem. This is due to the fact that addition of new components in the calculated model constantly results in a reduction of the residual error. In a paper by Pedersen *et al*. [P3] a JackKnife [68] procedure is presented in order to enhance the selection of the correct number of components to fit. Figure 13 shows a diagram of the proposed JackKnife procedure.

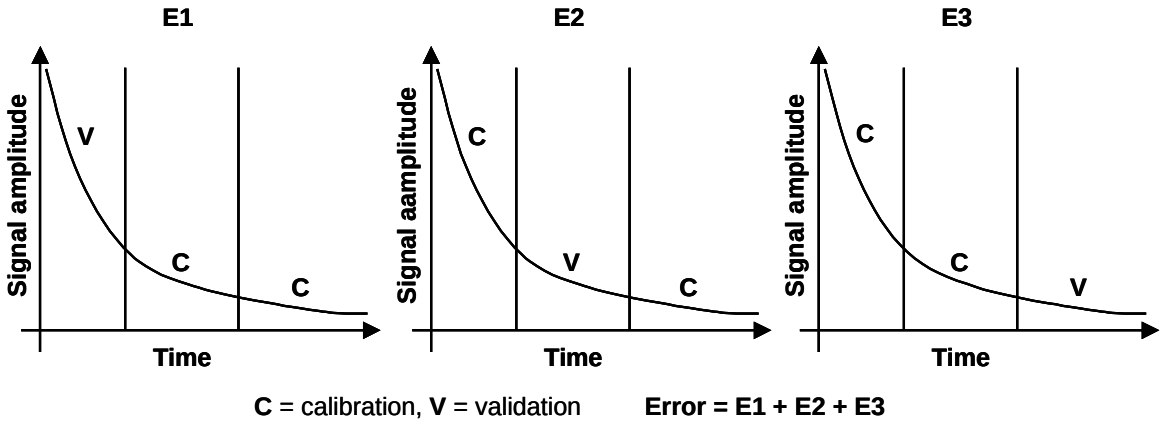

Figure 13 Diagram of the proposed JackKnife procedure in the situation where three segments are used. The three models calculated are displayed.

Thus by excluding the initial part of the relaxation signal the rank of the remaining relaxation profile is decreased. Therefore, the remaining signal is exponentially fitted and the amplitudes and time constants used to model the excluded part, the residual will significantly increase as a consequence and it can be concluded that too few components have been used. It has been found that the procedure stabilises the selection of components in the exponential fit procedure.

4 CURRENT AND FUTURE TRENDS

4.1 DATA ACQUISITION

4.1.1 From univariate to multivariate

A look at food science literature reveals that a lot of work has been published on the use of LF-NMR for determining different quality parameters in meat [69,70,71,72,73,74]. In particular, fat and water content have been widely studied [49,58,71,75].

Traditionally, fat content has been determined in meat samples by drying the samples overnight at 105°C and subsequently using a Hahn spin echo. Then the echo amplitude is related directly to the fat content of the sample.

Pedersen *et al.* [P6] suggest a way to enhance the performance of an existing method for measuring fat in meat in raw samples without drying the samples prior to measuring. The idea is based on the fact that the existing method can be extended from a univariate to a multivariate approach and takes advantage of the multivariate data analytical tools available.

A variety of pulse experiments may be suggested for measuring the fat content in meat. Some of the most obvious are the Hahn spin echo and the CPMG pulse experiment, but methods based on pulsed field gradients (PFG) for suppression of the water signal can also be used. The most common PFG experiments are the gradient stimulated echo and gradient spin echo experiments [76,77].

Pulsed field gradients are commonly used for suppression of water in food samples for enhancing the signal of some other component, but traditional application includes determination of diffusion constants of the different components in a sample [78,79,80] for determination of droplet size distribution or pore distribution in various samples [81]. Furthermore, it is field gradients that enables imaging by controlling the spatial acquisition which is required in order to produce an image [82,83,84,85].

Just as the CPMG experiment can be considered an extension of the Hahn spin echo experiment, enabling acquisition of entire relaxation curves, the PFG stimulated or spin echo experiment can also be extended by appending a 180-degree pulse train (as seen in CPMG) following the gradient echo. The

result is that a univariate pulse experiment has been converted into a multivariate pulse experiment and the range of data analytical tools has been greatly expanded.

The intention of the study by Pedersen *et al.* [P6] was to investigate univariate versus multivariate as well as no gradients versus application of gradients. Therefore, a Hahn spin echo, a CPMG, a PFG stimulated echo and a PFG stimulated echo extended as describe above were measured for a series of minced meat samples before and after drying. The following table shows the performance of the different pulse experiments for prediction of the fat content in the samples.

Tabel 1 Models for prediction of fat content in meat by LF-NMR calculated for the different pulse experiments for untreated and dried samples. *r* is the correlation, RMSECV is the root mean square error of cross validation and #PC indicates the number of components used in the different models.

Treatment	Pulse experiment	*r*	RMSECV	# PC
Untreated	SPIN-ECHO	0.61	2.93	-
	CPMG	0.98	0.76	5
	PFG-ECHO	0.91 (0.92)	1.50 (1.38)	-
	PFG-CPMG	0.99	0.49	2
Dry	SPIN-ECHO	1.00	0.25	-
	CPMG	1.00 (1.00)	0.26 (0.21)	1
	PFG-ECHO	1.00	0.24	-
	PFG-CPMG	1.00	0.28	2

Table adapted from paper. The models in brackets correspond to removal of two samples appearing to be outliers.

The first result that catches to the eye is the tremendous improvement in the prediction error for untreated samples simply by moving from the univariate spin echo to the multivariate CPMG experiment. Comparing the untreated samples and univariate models (SPIN-ECHO and PFG-ECHO) it is seen that application of a gradient for water suppression also significantly reduces the prediction error. The best performance on untreated samples is seen when the multivariate gradient pulse experiment is used (PFG-CPMG).

The results obtained for the dried samples clearly show that there is nothing gained by introducing either gradients or multivariate data modelling. In the case of dried samples only one factor varies in the samples and only minor

interfering signals may be present. Thus, the situation is simple enough for the univariate linear regression to handle.

Comparison of the traditional approach (spin echo on dry samples) with the best performing method for untreated samples shows, that for optimal predictive performance the traditional method on dried samples is still substantially better. However, comparing the result of the different models calculated for untreated samples an impressing improvement has been achieved using the multivariate pulsed field gradient pulse experiment and the prediction error is approaching that of the traditional method. Therefore, if the parameters for the PFG-CPMG method were optimised, it would be expected that the prediction error for untreated samples could be even further reduced. Secondly, in some applications a prediction error of 0.49 % fat may be acceptable taking into account the fact that the lengthy procedure of drying the samples has been eliminated.

In discussions following the publication of this work it has been suggested that it might have been a better choice to use a PFG spin echo experiment rather than a PFG stimulated echo experiment for the extension into a multivariate experiment. This is due to the fact that it is mathematically difficult to predict the resulting signal when stimulated echoes and spin echoes are mixed as seen in the applied set-up. However, at the time the work was carried out at the Swedish Meats R&D Institute a PFG stimulated echo experiment was routinely used for fat determination in meat. It was therefore decided to base the multivariate extension of a gradient experiment on this method, since appropriate experimental parameters were known and could be used directly (only tau and the number of echoes needed to be chosen).

It has clearly been demonstrated in this work that there are a number of new possibilities for improving existing methods as well as developing new methods for enhancing the performance of LF-NMR. Apparently, one such method is simply to rewrite pulse experiments in a way that enables acquisition of multivariate data and analyse these new data with multivariate data analytical techniques.

4.1.2 From multivariate to multi-way

It has previously been described that it is possible to acquire T_1-weighted CPMG data in an experiment where a CPMG pulse experiment is applied within the INVREC experiment following the recovery delay. The obtained data structure can be plotted as a 2D landscape, and if a number of such landscapes were stacked to generate a cuboid, it would be obvious to attempt application of three-way data analysis for optimal handling of such data.

One of the previous treatments of this kind of data was performed by Rutledge and Barros [47] where they examined true T_1-weighted CPMG data by means of ANOVA on the unfolded data and subsequently performed PLS on the unfolded data. Although they did not apply multi-way data analytical methods, they did observe irregularities in the refolded 2D variance landscapes resulting from the ANOVA that could not be accounted for. These irregularities appeared to occur for inversion values close to the inversion point where the sign of the signal changed from negative to positive. Closer examination of data acquired in this region has shown apparent non-exponential decay behaviour, which might simply originate from summing positive and negative exponentials or relate to a problem in data due to the significantly decreased signal-to-noise ratio.

Unpublished work performed in our own lab with the same kind of data applying three-way methods revealed that apparently data are not trilinear and a Parallel Factor (PARAFAC) model (see section 4.2 below) could not be successfully calculated on the basis of these data. This result is somewhat surprising, but one possible explanation might be that the PARAFAC model requires as equal number of components in the second and third mode and there is no guarantee that this requirement is actually fulfilled. If there are different numbers of T_1 and T_2 components present in the samples, a PARAFAC2 model should be used instead, since in this model different numbers of components may be specified for the T_1 and T_2 modes. Application of PARAFAC2 to this kind of data still remains to be tested.

Since T_1-weighted CPMG data might not be trilinear, different additional possibilities for generating trilinear data have been considered and to some degree tested. Completely new pulse experiments have been considered since it is not only through T_1-weighting that 2D data can be obtained. A

modification of the CPMG pulse experiment where tau remains constant while – through some initial steps – the signal is manipulated to produce a T_2-weighting of the acquired decay profiles is possible. Furthermore, it is possible to achieve T_1- and T_2-weighting through application of different values of the relaxation delay shorter than the appropriate value of the sample or different values of tau. This method is used for contrast purposes in magnetic resonance imaging (MRI).

Work is currently in progress to test these different possibilities for acquisition of true trilinear data with LF-NMR.

4.2 DATA ANALYSIS – MULTI-WAY

For a number of years multi-way analysis has been applied in a variety of research areas, and a large number of algorithms are available for this kind of analysis. One of the original formulations of a solution to this kind of problem is the Parallel Factor (PARAFAC) model proposed by Harshman in 1970 [86] and later elaborated upon [87,88].

Trilinear data follow the model:

$$x_{ijk} = \sum_{n=1}^{N} a_{in} b_{jn} c_{kn} + e_{ijk}, \; i = 1,..,I; \; j = 1,..,J; k = 1,..,K \qquad \text{Eq. 11}$$

The data are held in the elements x_{ijk}, the parameter a_{in} holds the scores pertaining to the first mode typically gathered in **A** ($I \times N$), b_{jn} holds the loadings pertaining to the second mode held in **B** ($J \times N$), c_{jn} holds the loadings pertaining to the third mode held in **C** ($K \times N$) and e_{ijk} holds residual unexplained variation. The elements x_{ijk} are held in a three-way array of dimension $I \times J \times K$ and are triply subscripted, meaning that the data can be arranged in a three-way box of data as opposed to ordinary doubly subscripted data corresponding to a matrix.

In bilinear modelling of matrices (e.g. PCA) the parameters are only identified up to rotation unless constraints such as non-negativity are applied. Thus, even though a set of several LF-NMR profiles follow the bilinear model, it is not possible to actually find the exponential profiles and their corresponding amplitudes, because an infinity of solutions provide the same fit. For the PARAFAC model, however, the parameters are uniquely identified up to trivial scaling and permutation. Hence, if the data follow the

model, the individual components can be identified directly [88]. Until recently, application of multi-way methods to LF-NMR data has not be possible since apparently true trilinear data has not yet been generated, as discussed in the previous section.

In 1997, however, an idea was conceived by Windig & Antalek [89] whereby 2D data can be created from 1D data containing underlying exponential components. This is achieved by using the fact that in an exponential curve the time constant is present in all points, however in smaller and smaller quantities, as displayed in Figure 14.

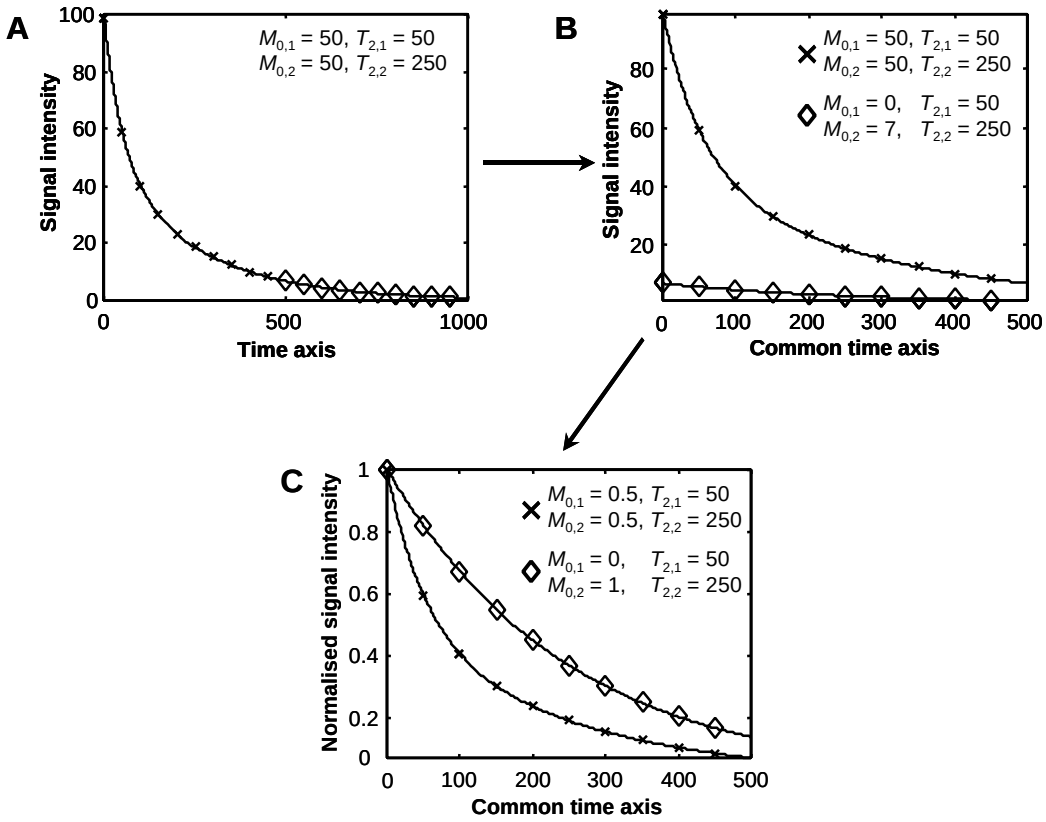

Figure 14 Illustration of the nature of multi-component exponential profiles, where all components theoretically contribute throughout the acquisition.

Here a two-component constructed relaxation curve has been generated and plotted (A). The curve can be split into two segments of equal size, and since the points in the curve are acquired on an equidistant time axis, the two segments can be plotted on a common axis (B). By normalising the signal intensity it becomes obvious that the two segments are not overlapping and the two segments make a bilinear matrix (C). This is supported by the fact

that bi-exponential fit of the two segments gives the same time constants, but that the amplitudes of the two components differ in the two segments.

This rearrangement of exponential data was proposed as a fast alternative to the trilinear least squares solution and was called Direct Exponential Curve Resolution Algorithm (DECRA). Similar ideas have also been applied in telecommunication under the generic name ESPRIT [90,91,92]. In their first application, Windig and Antalek applied DECRA to perform exponential curve resolution of first-order reaction kinetics ($C=C_0 \cdot e^{-k \cdot t}$) monitored by high-resolution NMR. Since the first publication of DECRA, several authors have applied the algorithm or modifications of it in a series of applications such as multivariate image analysis based on magnetic resonance images [93,94,95,96] short-wavelength near infrared analysis [97], UV-VIS [98,99] and solid state NMR and mid-infrared [100]. Common to these applications is the step of generating the three-way data array based on a measured two-way matrix, whereas different trilinear algorithms are used to perform the deconvolution.

None of the previously performed work has been performed on LF-NMR. However, Pedersen *et al.* [P5] have shown that the general concept of the approach can also be applied to LF-NMR, and in a later paper by Pedersen *et al.* [P7] an algorithm is presented to take care of the additional problems that must be taken into consideration when applying this method to LF-NMR data. This algorithm designed for handling LF-NMR data has been named SLICING, in acknowledgement of the fact that the 2D data are “sliced” in order to generate the 3D data, as will be described in the following.

If transverse LF-NMR data from two or more samples measured on an equidistant time axis can be approximated by:

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P}_{exp}^{T} \qquad \text{Eq. 12}$$

where $\mathbf{P}_{exp}$ contain N underlying profiles of length J ($J \times N$) which are distinct mono-exponentials, then it can be shown that the data can be rearranged into a three-way array $\underline{\mathbf{Y}}$ of size $I \times M \times K$ as depicted in Figure 15.

If $\mathbf{X}$ has elements x_{ij}, (i=1,..,I; j=1,..,J), then the three-way array $\underline{\mathbf{Y}}$ can consist of two sub-matrices of dimension $I \times J$-1, where the first sub-matrix contains the *first* J-1 columns of $\mathbf{X}$ and the second sub-matrix contains the *last* J-1

columns of **X**. Hence, the major part of the two matrices will be identical, but shifted "horizontally" by a fixed number of variables. The sub-matrices in the resulting three-way array will be referred to as *slabs* and the number of columns to shift between the two *slabs* (in this case one) will be referred to as *lag*.

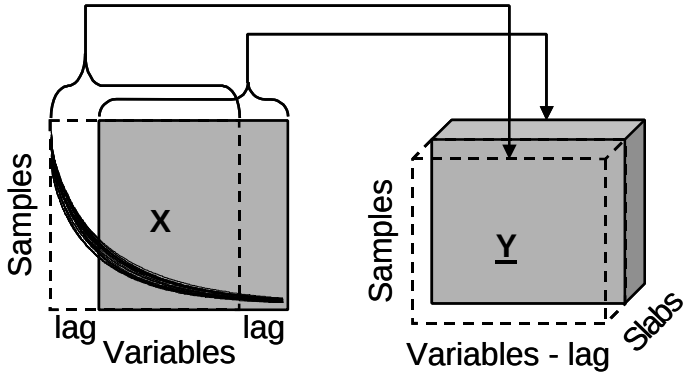

Figure 15 Illustration of the concept of rearranging a two-way data matrix into sub-matrices and placing them behind each other to create the three-way data structure. The figure shows the case where data have been lagged *lag* variables and two *slabs* are created.

In the original formulation of the DECRA algorithm, two *slabs* were generated using a *lag* of one variable and the trilinear deconvolution was based on the Generalised Rank Annihilation Methods (GRAM). In practice, there is nothing to hinder *lagging* the data more than one variable, still maintaining two *slabs* [98] or leading to possible generation of three or more *slabs* in the third mode [99]. When the dimension of the third mode increases, the three-way array $\underline{\mathbf{Y}}$, with elements y_{imk}, will now have the dimension i=1,..,I; m=1,..,J-L·(K-1); k=1,..,K with $2 \leq K \leq L+1$ where L is the selected *lag* and K is the specified number of *slabs*. With the introduction of more than two *slabs*, there are now three meta-parameters to optimise, i.e., the number of variables to *lag*, the number of *slabs* to generate and the number of factors to resolve, N. When the dimension of the third mode increases, GRAM can no longer be used and Direct Trilinear Decomposition (DTLD) is applied instead. In the situation where *lag* is one and *slabs* are two, GRAM and DTLD return the same solution.

In the work by Pedersen *et al*. [P7] the influence of a number of different parameters is tested on simulated data, comparing different standard set-ups

for values of *lag* and *slab*. Besides the proposed SLICING algorithm, a different approach for generating the three-way structure utilising the large number of variables typically seen in LF-NMR relaxation data was also tested. If in this alternative method it is still assumed that *K* describes the number of *slabs* to generate, then every *K*'th variable starting with the first variable describes the first *slab*, every *K*'th variable starting with the second variable will describe the second *slab* and so forth. The advantage of this approach should be that there is no redundancy in data and the structure in the residual commonly seen in noise in the normal SLICING approach is therefore eliminated.

In a series of initial studies however, somewhat surprisingly it has turned out that ordinary discrete exponential fitting performed on the entire data matrix appears to perform better than any of the SLICING set-ups tested. The iterative nature of the algorithm used for exponential fitting as well as the sensitivity to appropriate start values was expected to be a weakness, but this turns out not to be the case as long as good start guesses are used. Work is being continued with the SLICING approach in order to get at better understanding of the effects that influence the performance of the algorithm and currently the influence of a number of different parameters are being tested in an experimental design. SLICING may possibly see its best application as a means of estimating start guesses for the exponential matrix fit, thus speeding up this algorithm.

In elaborating on the SLICING algorithm Pedersen *et al.* [P7] suggest a new algorithm for performing fast and easy phase rotation of time domain LF-NMR quadrature data. Phase rotation is an important topic when performing curve fitting on the acquired LF-NMR relaxation data, since magnitude correction of quadrature data will enforce a non-exponential bend in data when the signal intensity approaches zero.

The development of phase rotation algorithms has previously focused mainly on high-field NMR data where a great deal of attention and effort has been dedicated to solving this problem [101].

The fundamental idea of the proposed algorithm, which is called Principal Phase Correction (PPC), is similar to that of PCA, where the variability in the

data set is maximised in an orthogonal subspace. Thus, calculating a two component PCA model using the raw quadrature data as input will result in the phase rotated data residing as the scores of the first principal component. This may at first glance seem somewhat confusing, but the purpose of the phase rotation algorithm is to maximise the signal in one channel. The computation is quite simple and can easily be performed using singular value decomposition (SVD).

4.3 HARDWARE – ONE-SIDED MAGNET LAYOUT

Despite various advantages, such as being non-invasive and non-destructive, low-field bench-top NMR instruments still have a number of drawbacks. And in a lot of applications the non-invasive and non-destructive description does not hold, since most instruments have small-bore probe-heads and small magnet gaps, and sub sampling is often necessary. Furthermore, the instruments require stable ambient temperature and vibrations as well as extraneous magnetic fields should be eliminated. Finally, due to the weight of the permanent magnets used, the instruments cannot be moved around easily. All these factors result in LF-NMR being a method that is mainly used in a laboratory environment and on- and at-line applications consequently still await a breakthrough.

Some applications of portable NMR devices have previously been shown, for instance in well logging in the oil industry where an NMR device is lowered into the bore hole to assist in estimating the potential for profitable utilisation of the current well [102]. A device carried on the back of a tractor for measuring the moisture content of the soil on a field has also been developed. Although portable, these instruments can hardly be considered handy or suitable for implementation in a production line, but they do set an example for inspiration.

In an attempt to build an instrument that would not be limited by sample geometry and at the same time be light and portable, a new hardware layout was developed, resulting in a hand-size NMR probe called MObile Universal Surface Explorer (MOUSE). The MOUSE is a small LF-NMR device that was first proposed in 1996 by Eidmann *et al.* [6] and further theoretically

described by Blümich *et al.* [103] and Bãlibanu *et al.* [104]. A drawing of its one-sided magnet layout is shown in Figure 16.

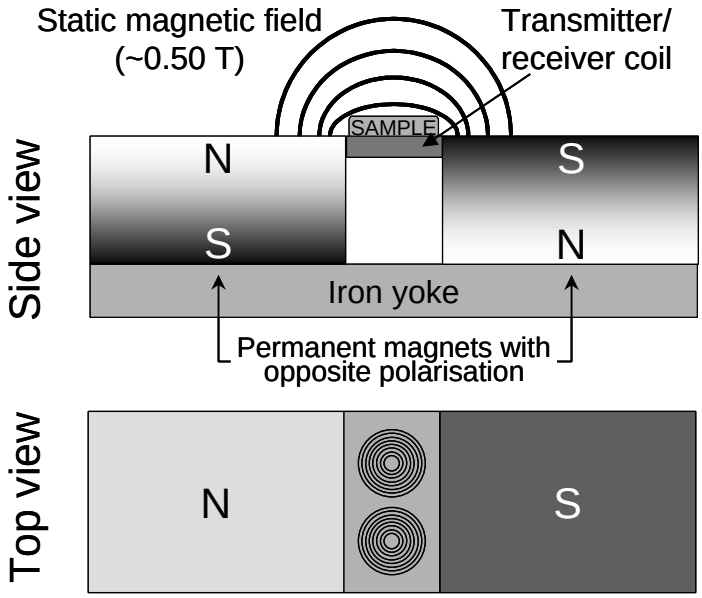

Figure 16 Drawing of the NMR-MOUSE both as side and top view. Magnetic filed lines are pictured.

While sample geometry is unrestricted with the one-sided magnet layout of the MOUSE, there is still a price to be paid for this improvement by severely reduced magnetic field homogeneity introduced by a large B_0-gradient. Since the field strength rapidly decreases as a function of distance from the surface of the MOUSE, measurements are only possible close to the surface of the probe and consequently restricted to the surface of the sample. The strong gradient in the magnetic field (in the order of 10 T/m) reduces the sample volume measured by the MOUSE into a surface area through the sample rather than a volume of sample, which is normally the case for standard magnet layout. The increased inhomogeneity results in a significant reduction of the signal to noise ratio in the acquired data when compared to a normal LF-NMR instrument.

Most experiments performed to date using the MOUSE have focused on polymer materials like rubber [105,106,106] looking at different properties such as weathering [107] and cross-link density [108], but a study of anisotropy in tendon has also been performed [109]. In a more general paper application of the MOUSE to soft matters is discussed [110].

There are a number of reasons why polymers have been an obvious choice of material to investigate using the MOUSE. Among others, these relate to the fact that T_1 is fairly short for most polymers, allowing application of multiple scans in order to enhance signal-to-noise, while T_2 are of appreciable length, allowing the signal to be detected. Furthermore, most polymers are quite homogeneous and due to the restricted mobility relaxation is hardly affected by the strong magnetic field gradient. Diffusion of mobile components, however, is a significant problem to consider when applying the MOUSE to materials with high water content such as foodstuffs, since the strong magnetic field gradient of the MOUSE will strongly influence the relaxation rate of the mobile protons. In the case of water, the resulting effect is that the water signal rapidly decays to zero and the transverse relaxation time constant is significantly decreased. This effect is clearly visible from Figure 17 where signals from water and oil acquired using a normal LF-NMR instrument as well as the MOUSE are plotted.

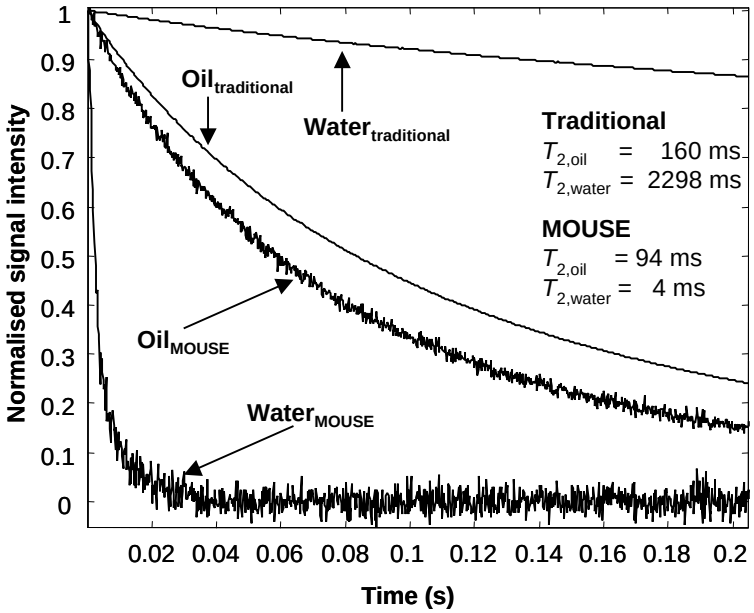

Figure 17 Relaxation profiles and corresponding time constants for pure water and oil acquired by the normal LF-NMR instrument ("traditional") and the NMR-MOUSE

It is interesting to observe that the position of the relaxation profile for water in relation to oil shifts when comparing the acquisitions from the two instruments. Furthermore, examination of the time constants reported in the two plots also indicates the strong influence on the water relaxation exerted by the magnetic gradient field.

Pedersen *et al.* [P8] applied the NMR-MOUSE to a series of oil-in-water emulsions with different relative contents and the predictive performance was evaluated and compared to that of a normal LF-NMR instrument which is known to be able to predict the oil content in emulsions [19,36,102,111].

In Figure 18 the data acquired by the normal LF-NMR instrument and the NMR-MOUSE can be seen. Two things are apparent when comparing the two plots: the signal to noise level (S/N) is significantly lower for the MOUSE and, as expected, the order of the profiles in the two plots is reversed due to the strong magnetic field gradient of the MOUSE.

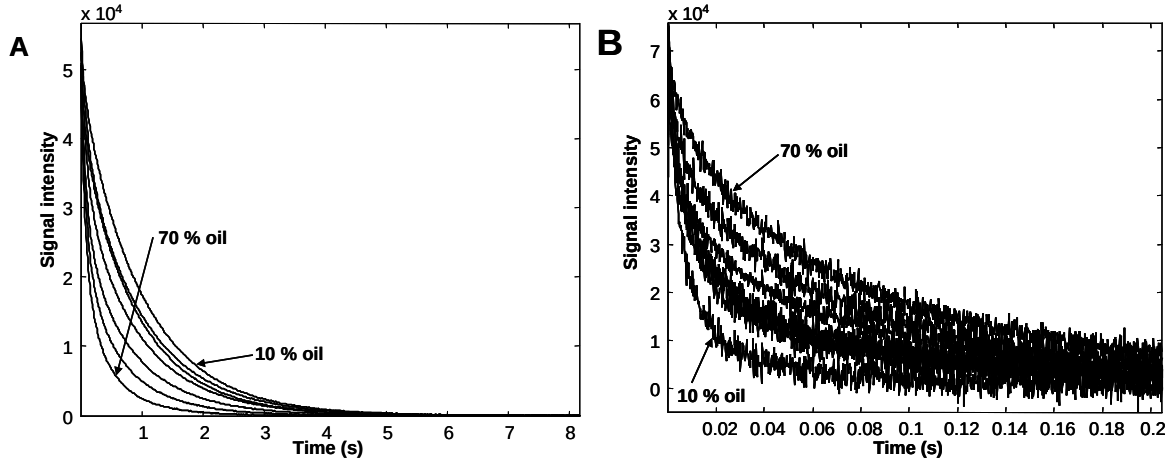

Figure 18 Plot of the acquired relaxation profiles from seven oil-in-water emulsions measured on the normal LF-NMR instrument (A) and the MOUSE (B)

Data analysis on the normal LF-NMR data is performed by means of traditional bi-exponential fitting followed by linear regression of the ratio calculated between the amplitudes as well as PLS performed on the full relaxation decays. Data acquired by the MOUSE is also analysed by these two methods, and, in addition, MLR is performed on the two amplitudes and two time constants from the bi-exponential fit. The resulting models can be seen in the Tabel 2.

For the normal LF-NMR data quite good performance is seen for both the linear regression and PLS models. The prediction error, however, is significantly smaller for the PLS model, which is also the case with bias, a measure of the average difference between measured and predicted values.

Tabel 2 Results of the different models applied for the acquired data both for the MOUSE and the traditional LF-NMR instrument. *r* is the correlation, RMSECV is the root mean square error of cross validation and bias is a measure of the average difference between measured and predicted values.

	Data analysis applied	*r*	RMSECV	bias
LF-NMR	Linear regression	1.00	2.21	-1.14
	PLS	1.00	1.37	0.29
MOUSE	Linear regression	0.99	5.55	4.09
	MLR	0.98	3.79	-0.29
	PLS	0.97	4.44	0.24

Table adapted from paper. The models in brackets correspond to removal of two samples appearing to be outliers.

For the MOUSE data it is observed that the correlation is very good for the linear regression, but that both the prediction error and bias are rather high. Here the advantage of multivariate modelling is even more pronounced with the surprise that the MLR model based on the amplitudes and time constants from a bi-exponential fit produces the best model in terms of prediction error. The reason for this is not known, but might be due to the fact that the data set contains a rather small number of samples. The high noise level might also present a problem, although PLS is normally considered to be very robust to noisy data.

It is concluded from the study that with application of multivariate data analysis the NMR-MOUSE can in fact be used to analyse food samples with high water content, despite the influence of the strong field gradient. Further work is required in order to properly characterise the performance of the MOUSE on other samples with high water content, but the obtained result is very interesting and promising and opens doors to many applications of small portable NMR instruments. New applications and improved MOUSE designs are being tested, such as a sweep MOUSE for 1D profiling, which can also simply be used for increasing the depth of acquisition for normal relaxation measurements. Furthermore an imaging MOUSE is being developed for surface imaging. The improved depth sensitivity means that the MOUSE can be used for applications other than those just surface measurements. From the performed work it seems clear that chemometrics will have to be a natural part of these applications in order for the device to produce acceptable results.

4.4 DYNAMICS

One of the great advantages of LF-NMR compared to most other spectroscopic methods is the ability to accurately control sample temperature while continuously measuring. This enables a wide range of possibilities for manipulating the sample to make dynamic studies of processes occurring in food samples. This kind of analysis has become increasingly popular in recent years and has, for instance, been used to study the changes taking place during the process of baking bread with respect to the hydration of starch [P4]. Viscoelastic properties of food materials have been studied by rheology using MRI looking at shear-rate dependent viscosity such as shear thinning, shear thickening, apparent slip and yield stress [112,113]. MRI has also been used to study liquid triglyceride migration in chocolate at different temperatures as well as in interface layers between dark chocolate and hazelnut [114]. Furthermore, the digestion process in the human stomach followed by emptying to the duodenum has been studied using MRI. In this study, the movements of the stomach, the separation of the food in the stomach as well as the rate of emptying was examined. The list of possible applications and applications already tested is long.

It is expected that dynamic studies will play a more and more dominant role in studies of processes taking place in food samples and thereby allow researchers to take a leap forward in understanding the mechanisms controlling food process and quality. This will eventually lead to higher quality products, for example, products with more uniform quality or products with better shelf life due to increased understanding of the properties of water in food, all of which is hoped to benefit the consumers.

5 CONCLUSIONS AND PERSPECTIVES

From the previous chapters it should be apparent that feasibility of application of LF-NMR in food science, and even more so in process control and analysis, has been greatly improved over the years. This advancement is partly due to improvements in hardware, but from the literature studied it seems that the largest improvement has come from the introduction of multivariate data analysis.

There exists a large range of possibilities for implementing LF-NMR in food analysis, which cover a number of different research areas such as the analysis of fish, oil seeds, bread, meat, emulsions and meat. It is believed that revision of existing methods in most applications will result in greatly improved performance, and possibly facilitate new applications. Such applications could be the analysis of components present in small amounts, in data with a high noise level or in data with closely spaced components.

By rewriting the normal pulsed field gradient stimulated echo experiment into a pulsed field gradient CPMG pulse experiment in order to acquire multivariate rather than univariate data a significant improvement in the prediction of fat in meat has been achieved. The changed pulse experiment enables multivariate data analysis, which offers the advantage of outlier detection as well as handling of extremely collinear data. The prediction error has been reduced from 1.50 % fat to 0.49 % fat (range from 1.2 % to 15 % fat) for measurement on raw minced meat samples. This should be compared to the result obtained by the traditional method where the samples are dried prior to the measurement, resulting in a prediction error of 0.25 % fat. This error is still much better, however, the fact that the fat content can be quite accurately measured without drying means a tremendous reduction in the time of the analysis. This is very important, since the time of analysis is crucial in many applications such as process or quality control. It appears that such an approach to improve the prediction error has never been applied before.

As a consequence of the problem of drifting time constants between samples when exponentially fitting is performed, a new algorithmic approach has been elaborated upon and tested in an attempt to establish a better way of calculating the underlying time constants in a series of samples. The method

tested has been called SLICING and utilises the fact that a pseudo three-way structure can be generated for data with underlying exponential components. The method applies three-way analysis for unique deconvolution of the latent factors. It seems, however, that a much simpler approach based on simultaneous exponential fitting of a collection of samples, rather than fitting one sample at the time, results in a more robust and even better deconvolution. A combination of the two different approaches is under consideration for optimal speed and performance.

No doubt, there are a number of additional possibilities in data analysis that have not yet been exploited. It appears necessary to think more in terms of a complete solution where data analysis and data acquisition go hand in hand. The idea is that data acquisition may be revised in order to produce the specific kind of data required by a given algorithm, or the other way around, that more work is put into development of better algorithms designed for utilising new techniques and ideas. This concept will require new thinking and, more so, it will require acceptance of the fact that new methods based on soft models do in fact provide trustworthy results.

The presented work performed on the NMR-MOUSE is the first of its kind where this small portable LF-NMR instrument has been applied to food-like samples with highly mobile components. The advantage of the MOUSE compare to conventional LF-NMR instruments is that the one-sided layout allows measurement on large samples, being completely non-invasive and non-destructive. Furthermore, the MOUSE is a small handheld device, making it more appropriate for a number of applications. The readily diffusing mobile components cause problems in the inhomogeneous magnetic field of the MOUSE, since the signal rapidly decays. This, however, is nicely handled by the chemometric techniques applied, reducing the prediction error from 5.55 % oil with the traditional data handling procedures to 3.79 % oil with a multiple linear regression on a series of oil in water emulsions (from 10 to 70 % oil).

Other applications of portable NMR devices have previously been shown, for instance, in well logging in the oil industry and a device carried on the back of a tractor for measuring the moisture content of the soil on a field has also been developed.

New applications and improved MOUSE designs are being tested, such as a sweep MOUSE for 1D profiling or simply improving the depth of acquisition, and an imaging MOUSE is being developed. The improved depth sensitivity means that the MOUSE can be used for applications other than those just surface measurements. From the performed work it seems clear that chemometrics will have to be a natural part of these applications in order for the device to produce acceptable results.

New designs either focused on general applications or devices designed for specific purposes clearly have great possibilities for boosting the number of industrial applications currently seen.

Despite the obvious advantages of multivariate data analysis that have been shown in this work, only a limited number of researchers in the LF-NMR area appear to have “seen the light”. It is hoped, though, that a broader range of researchers and industrial people will start applying chemometrics, resulting in “better value for money” and a wider range of applications. This has been seen in recent years, particularly in the food industry where the interest for using chemometrics in combination with other spectroscopic methods has increased, leading to better control and understanding of processes and faster analysis of a variety of food materials.

6 REFERENCE LIST

1. Gorter, C.J.: Negative resluts of an attempt to direct nuclear magnetic spins, *Physica*, 1936, **3**, 995-998

2. Bloch, F.; Hansen, W.W.; Packard, M.: Nuclear Induction, *Physical Review*, 1946, **69**, 127

3. Purcell, E.M.; Torrey, H.C.; Pound, R.V.: Resonance absorption by nuclear magnetic moments in a solid, *Physical Review*, 1946, **69**, 37-38

4. Proctor, W.G.; Yu, F.C.: The dependence of a nuclear magnetic resonance frequency upon chemical compounds, *Physical Review*, 1950, **77**, 717-717

5. Claridge, T.D.W.: "High-resolution NMR techniques in organic chemistry", Pergamon, Elsevier Science Ltd., Oxford, 1999

6. Eidmann, G.; Savelsberg, R.; Blümler, P.; Blümich, B.: The NMR MOUSE, a mobile universal surface explorer, *Journal of Magnetic Resonance, Series A*, 1996, **122**, 104-109

7. Hahn, E.L.: Spin echoes, *Physical Review*, 1950, **80**, 580-594

8. Carr, H.Y.; Purcell, E.M.: Effects of diffusion on free precession in nuclear magnetic resonance experiments, *Physical Review*, 1954, **94**, 630-638

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

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

11. Brøndum, J.; Byrne, D.V.; Bak, L.S.; Bertelsen, G.; Engelsen, S.B.: Warmed-over flavour in porcine meat - a combined spectroscopic, sensory and chemometric study, *Meat Science*, 2000, **54**, 83-95

12. Brøndum, J.; Munck, L.; Henckel, P.; Karlsson, A.; Tornberg, E.; Engelsen, S.B.: Prediction of water-holding capacity and composition of porcine meat by comparative spectroscopy, *Meat Science*, 2000, **55**, 177-185

13. Samuelsson, E.G.; Vikelsøe, J.: Estimation of the amount of liquid fat in cream and butter by low resolution NMR, *Milchwissenschaft*, 1971, **26**, 621-625

14. Van den Enden, J.C.; Rossell, J.B.; Vermaas, L.F.; Waddington, D.: Determination of the solid fat content of hard confectionery butters, *Journal of the American Oil Chemists' Society*, 1982, **59**, 433-439

15. Gribnau, M.C.M.: Determination of solid/liquid ratios of fats and oils by low-resolution pulsed NMR, *Trends in Food Science & Technology*, 1992, **3**, 186-190

16. Tiwari, P.N.; Gambhir, P.N.; Rajan, T.S.: Rapid and non-destructive determination of seed oil by pulsed nuclear magnetic resonance technique, *Journal of American Oilchemists' Society*, 1974, **51**, 104-109

17. Gambhir, P.N.: Applications of low-resolution pulsed NMR to the determination of oil and moisture in oilseeds, *Trends in Food Science & Technology*, 1992, **3**, 191-196

18. Brosio, E.; Conti, F.; Di Nola, A.; Scorano, O.; Balestrier, F.: Simultaneous determination of oil and water content in olive husk by pulsed low resolution nuclear magnetic resonance, *Journal of Food Technology*, 1981, **16**, 629-636

19. Brosio, E.; Conti, F.; Di Nola, A.; Scalzo, M.; Zulli, E.: Oil and water determination in emulsions by pulsed low-resolution NMR, *Journal of the American Oil Chemists' Society*, 1982, **59**, 59-61

20. Rubel, G.: Simultaneous determination of oil and water contents in different oilseeds by pulsed nuclear magnetic resonance, *Journal of the American Oil Chemists' Society*, 1994, **71**, 1057-1062

21. Wing, K.W.; Stroshine, R.L.; Krutz, G.W.: A modified Hahn echo pulse sequence for proton magnetic resonance (1H-MR) measurements of percent soluble solids of fruits, *Transactions of the American Society of Agricultural Engineers (ASAE)*, 1995, **38**, 849-855

22. Tellier, C.; Guillou, M.; Grenier, P.; Le Botlan, D.: Monitoring alcoholic fermentation by low-resolution pulsed nuclear magnetic resonance, *Journal of Agricultural and Food Chemistry*, 1989, **37**, 988-991

23. Lambelet, P.; Berrocal, R.; Ducret, F.: Low resolution NMR spectroscopy: A tool to study protein denaturation: I. Application to diamagnetic whey proteins, *Journal of Dairy Research*, 1989, **56**, 211-222

24. Gambhir, P.N.; Chahal, S.S.; Nagarajan, S.; Tiwari, P.N.: Proton spin-lattice relaxation study of cell water in wheat (Triticum aestivum L.) and barley (Hordeum vulgare L.) leaves, *Indian Journal of Experimental Biology*, 1991, **29**, 652-655

25. Monteiro-Marques, J.P.; Le Loch, C.; Wolff, E.; Rutledge, D.N.: Monitoring freeze-drying by low resolution pulse NMR: Application to the determination of sublimation endpoint, *Journal of Food Science*, 1991, **56**, 1707-1710+1728

26. Lambelet, P.; Renevey, F.; Kaabi, C.; Raemy, A.: Low-field nuclear magnetic resonance relaxation study of stored or processed cod, *Journal of Agricultural and Food Chemistry*, 1995, **43**, 1462-1466

27. Steen, C.; Lambelet, P.: Texture changes in frozen cod mince measured by low-field nuclear magnetic resonance spectroscopy, *Journal of the Science of Food and Agriculture*, 2000, **75**, 268-272

28. Roudaut, G.; Van Dusschoten, D.; Van As, H.; Hemminga, M.A.; Le Meste, M.: Mobility of lipids in low moisture bread as studied by NMR, *Journal of Cereal Science*, 1998, **28**, 147-155

29. Ruan, R.R.; Long, Z.Z.; Song, A.J.; Chen, P.L.: Determination of the glass transition temperature of food polymers using low field NMR, *Journal of Food Science and Technology*, 1998, **31**, 516-521

30. Brosio, E.; Conti, F.; Di Nola, A.; Sykora, S.: Correlation between iodine number and proton relaxation times in maize oil, *Journal of Food Technology*, 1981, **16**, 67-72

31. El Khaloui, M.; Rutledge, D.N.; Ducauze, C.J.: Monitoring hydrogenation in the margarine industry by low resolution pulsed NMR, *Journal of the Science of Food and Agriculture*, 1990, **53**, 389-393

32. Provencher, S.W.; Dovi, V.G.: Direct analysis of continuous relaxation spectra, *Journal of Biochemical and Biophysical Methods*, 1979, **1**, 313-318

33. Lillford, P.J.; Clark, A.H.; Jones, D.V.: *Distribution of water in heterogeneous food and model systems*, in "Water in Polymers" (Rowland, R. J., eds.), pp. 177-194, American Chemical Society, Washington, D.C., 1980

34. Hills, B.P.; Remigereau, B.: NMR studies of changes in subcellular water compartmentation in parenchyma apple tissue during drying and freezing, *International Journal of Food Science and Technology*, 1997, **32**, 51-61

35. Le Botlan, D.; Rugraff, Y.; Martin, C.; Colonna, P.: Quantitative determination of bound water in wheat starch by time domain NMR spectroscopy, *Carbohydrate Research*, 1998, **308**, 29-36

36. Le Botlan, D.; Wennington, J.; Cheftel, J.C.: Study of the state of water and oil in frozen emulsions using time domain NMR, *Journal of Colloid and Interface Science*, 2000, **226**, 16-21

37. Bertram, H.C.; Andersen, H.J.; Karlsson, A.H.: Comparative study of low-field NMR relaxation measurements and two traditional methods in the determination of water holding capacity of pork, *Meat Science*, 2001, **57**, 125-132

38. Hansen, P.C.: Regularization Tools: A Matlab package for analysis and solution of discrete ill-posed problems, *Numerical Algorithms*, 1994, **6**, 1-35

39. Butler, J.P.; Reeds, J.A.; Dawson, S.V.: Estimating solutions of first kind integral equations with nonnegative constraints and optimal smoothing, *SIAM Journal of Numerical Analysis*, 1981, **18**, 381-397

40. Tellier, C.; Guillou-Charpin, M.; Le Botlan, D.; Pelissolo, F.: Analysis of low-resolution, low-field NMR relaxation data with the Padé-Laplace method, *Magnetic Resonance in Chemistry*, 1991, **29**, 164-167

41. Borgia, G.C.; Brown, R.J.S.; Fantazzini, P.: Uniform-Penalty inversion of multiexponential decay data, *Journal of Magnetic Resonance*, 1998, **132**, 65-77

42. Steinbrecher, G.; Scorei, R.; Cimpoiasu, V.M.; Petrisor, I.: Stable reconstruction of the T2 distribution by low-resolution NMR measurements and the classical Markov and Hausdorf momentum problem, *Journal of Magnetic Resonance*, 2000, **146**, 321-334

43. Davenel, A.; Marchal, P.: Discriminant analysis applied to moisture determination in raw materials for animal feed by pulsed NMR, *Transactions of the American Society of Agricultural Engineers (ASAE)*, 1992, **35**, 1891-1897

44. Tollner, E.W.; Hung, Y.C.: Magnetic resonance for measuring moisture in wheat, corn, soybean, pecans and peanuts, *Paper American Society of Agricultural Engineers*, 1990, **90-3008**, 1-22

45. Hernandez, C.V.; Rutledge, D.N.: Characterisation of cocoa masses: low resolution pulse NMR study of the effect of geographical origin and roasting on fluidification, *Food Chemistry*, 1994, **49**, 83-93

46. Rutledge, D.N.; Barros, A.S.; Gaudard, F.: ANOVA and factor analysis applied to time domain NMR signals, *Magnetic Resonance in Chemistry*, 1997, **35**, S13-S21

47. Rutledge, D.N.; Barros, A.S.: Method for detecting information in signals: application to two-dimensional time domain NMR data, *Analyst*, 1998, **123**, 551-559

48. Rutledge, D.N.; Barros, A.S.; Vackier, M.C.; Baumberger, S.; Lapierre, C.: *Analysis of time domain NMR and other signals*, in "Advances in Magnetic Resonance in Food Science" (Belton, P. S.; Hills, B. P.; Webb, G. A., eds.), pp. 203-216, The Royal Society of Chemistry, Cambridge, 1999

49. Brown, R.J.S.; Capozzi, F.; Cavani, C.; Cremonini, M.A.; Petracci, M.; Placucci, G.: Relationships between H-1 NMR relaxation data and some technological parameters of meat: A chemometric approach, *Journal of Magnetic Resonance*, 2000, **147**, 89-94

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

51. Wold, S.; Esbensen, K.; Geladi, P.: Principal component analysis - A tutorial, *Chemometrics and Intelligent Laboratory Systems*, 1987, **2**, 37-52

52. Thybo, A.K.; Bechmann, I.E.; Martens, M.; Engelsen, S.B.: Prediction of sensory texture of cooked potatoes using uniaxial compression, near infrared spectroscopy and low field ^{1}H NMR spectroscopy, *Lebensmittel-Wissenschaft und Technologie*, 2000, **33**, 103-111

53. Rutledge, D.N.; Barros, A.S.; Giangiacomo, R.: *Interpreting near infrared spectra of solutions by outher product analysis with time domain NMR*, in "Magnetic Resonance in Food Science - A view to the Future" (Webb, G. A.; Belton, P. S.; Gill, A. M.; Delgadillo, I., eds.), pp. 179-192, Royal Society of Chemistry, Cambridge, 2001

54. Wold, S.; Martens, H.; Wold, H.: *The multivariate calibration problem in chemistry solved by the PLS method*, in "Lecture notes in mathematics" (Ruhe, A.; Kågstrom, B., eds.), pp. 286-293, Springer Verlag, Heidelberg, 1983

55. Geladi, P.; Kowalski, B.R.: Partial least squares regression: A tutorial, *Analytical Chimica Acta*, 1986, **185**, 1-17

56. Martens, H.; Næs, T.: "Multivariate calibration", John Wiley & Sons, 1989

57. Thygesen, L.G.: PLS calibration of pulse NMR free induction decay for determining moisture content and basic density of softwood above fibre saturation, *Holzforschung*, 1996, **50**, 434-436

58. Gerbanowski, A.; Rutledge, D.N.; Feinberg, M.H.; Ducauze, C.J.: Multivariate regression applied to time domain nuclear magnetic resonance signals: Determination of moisture in meat products, *Sciences des Aliments*, 1997, **17**, 309-323

59. Airiau, C.; Gaudard, F.; Barros, A.S.; Rutledge, D.N.: Complexation stoichiometry determined by application of chemometrics to time domain nuclear magnetic resonance signals, *Analusis*, 1998, **26**, 70-74

60. Tornberg, E.; Wahlgren, M.; Brøndum, J.; Engelsen, S.B.: Pre-rigor conditions in beef under varying temperature- and pH-falls studied with rigometer, NMR and NIR, *Food Chemistry*, 2000, **69**, 407-418

61. Prony, R.: Essai Expérimental et Analytique sur les Lois de la Dilatibilité et sur celles de la Force Expansive de la Vapeur de l'Eau et de la Vapeur de l'Alkool, a Différentes Temperatures, *Journal de l'Ecole Polytechnique*, 1795, **1**, 24-76

62. Gardner, D.G.; Gardner, J.C.; Meinke, W.W.: Method for the analysis of multicomponent exponential decay curves, *The Journal of Chemical Physics*, 1959, **31**, 978-986

63. Marquardt, D.W.: An algorithm for least-squares estimation of non-linear parameters, *Journal of the Society for Industrial and Applied Mathematics*, 1963, **11**, 431-441

64. Schlesinger, J.: Fit to experimental data with exponential functions using the fast fourier transform, *Nuclear Instruments and Methods*, 1973, **106**, 503-508

65. Schreurs, S.; François, J.-P.: The combined use of the improved methods of Gardner and Prony for the qualitative and quantitative analysis of multicomponent decay curves, *Chemometrics and Intelligent Laboratory Systems*, 1998, **43**, 107-121

66. Wijnaendts van Resandt, R.W.; Vogel, R.H.; Provencher, S.W.: Double beam fluorescence spectrometer with subnanosecond resolution: Application to aqueous tryptophan, *Review of scientific instruments. New series*, 1982, **53**, 1392

67. Davenel, A.; Marchal, P.; Guillement, J.P.: *Rapid cooking control of cakes by low resolution NMR*, in "Magnetic Resonance in Food Science" (Delgadillo, I.; Gill, A. M.; Webb, G. A.; Belton, P. S., eds.), pp. 146-155, The Royal Society of Chemistry, Cambridge, 1995

68. Efron, B.: "The Jackknife, the Bootstrap and other resampling plans", Society for Industrial and Applied Mathematics, Philadelphia, PA, 1982

69. Wright, R.G.; Milward, R.C.; Coles, B.A.: Rapid protein analysis by low-resolution pulsed NMR, *Food Technology*, 1980, **34**, 47-52

70. Tipping, L.R.H.: The analysis of protein in fresh meats using pulsed NMR, *Meat Science*, 1982, **7**, 279-283

71. Renou, J.P.; Monin, G.; Sellier, P.: Nuclear magnetic resonance measurements on pork of various qualities, *Meat Science*, 1985, **15**, 225-233

72. Borisova, M.A.; Oreshkin, E.F.: On the water condition in pork meat, *Meat Science*, 1992, **31**, 257-265

73. Beauvallet, C.; Renou, J.P.: Applications of NMR spectroscopy in meat research, *Trends in Food Science & Technology*, 1992, **3**, 241-246

74. Tornberg, E.; Andersson, A.; Von Seth, G.: Water distribution in raw pork muscle (M. longissimus dorsi) of different meat qualities, *39th International Congress of Meat Science and Technology*, 1993,

75. Renou, J.P.; Briguet, A.; Gatellier, Ph.; Kopp, J.: Technical note: Determination of fat and water ratios in meat products by high resolution NMR at 19.6 MHz, *International Journal of Food Science and Technology*, 1987, **22**, 169-172

76. Stejskal, E.O.; Tanner, J.E.: Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient, *The Journal of Chemical Physics*, 1965, **42**, 288-291

77. Tanner, J.E.: Restricted self-diffusion of protons in colloidal systems by the pulsed gradient, spin-echo method, *The Journal of Chemical Physics*, 1968, **49**, 1768-1777

78. Callaghan, P.T.; Jolley, K.W.; Humphrey, R.S.: Diffusion of fat and water in cheese as studied by pulsed field gradient nuclear magnetic resonance, *Journal of Colloid and Interface Science*, 1983, **93**, 521-529

79. Cotts, R.M.; Hoch, M.J.R.; Sun, T.; Markert, J.T.: Pulsed field gradient stimulated echo methods of improved NMR diffusion measurements in heterogeneous systems, *Journal of Magnetic Resonance*, 1989, **83**, 252-266

80. Watanabe, H.; Fukuoka, M.: Measurement of moisture diffusion in foods using pulsed field gradient NMR, *Trends in Food Science & Technology*, 1992, **3**, 211-215

81. Fourel, I.; Guillement, J.P.; Le Botlan, D.: Determination of water droplet size distributions by low resolution PFG-NMR, *Journal of Colloid and Interface Science*, 1994, **164**, 48-53

82. Brosio, E.: Nuclear magnetic resonance: A multiparameter technique for in situ analysis, *Reviews in Analytical Chemistry*, 1995, **14**, 227-252

83. Belton, P.S.; Delgadillo, I.; Gil, A.M.; Webb, G.A.: "Magnetic resonance in food science", 1998

84. Hills, B.P.: "Magnetic resonance imaging in food science", John Wiley & Sons, Inc., New York, 1998

85. Popella, H.; Henneberger, G.: Design and optimization of the magnetic circuit of a mobile nuclear magnetic resonance device for magnetic resonance imaging, *COMPEL - The international journal for computation and mathematics in electrical and electronic engineering*, 2001, **20**, 269-278

86. Harshman, R.A.: Foundations of the PARAFAC procedure: Models and conditions for an 'explanatory' multi-modal factor analysis, *UCLA working papers in phonetics*, 1970, **16**, 1-84

87. Harshman, R.A.; Lundy, M.E.: PARAFAC: parallel factor analysis, *Computational Statistics and Data Analysis*, 1994, **18**, 39-72

88. Bro, R.: PARAFAC. Tutorial and applications, *Chemometrics and Intelligent Laboratory Systems*, 1997, **38**, 149-171

89. Windig, W.; Antalek, B.: Direct Exponential Curve Resolution Algorithm (DECRA): A novel application of the generalized rank annihilation method for a single spectral mixture data set with exponentially decaying contribution profiles, *Chemometrics and Intelligent Laboratory Systems*, 1997, **37**, 241-254

90. Roy, R.; Kailath, T.: Estimation of signal parameters via rotational invariance techniques, *IEEE ASSP magazine*, 1989, **37**, 984-995

91. Viberg, M.: Sensor array processing based on subspace fitting, *IEEE Transactions on Signal Processing*, 1991, **39**, 1110-1121

92. Sidiropoulos, N.D.; Bro, R.; Giannakis, G.B.: Parallel factor analysis in sensor array processing, *IEEE Transactions on Signal Processing*, 2000, **48**, 2377-2388

93. Windig, W.; Hornak, J.P.; Antalek, B.: Multivariate image analysis of magnetic resonance images with the direct exponential curve resolution algorithm (DECRA) - Part 1: Algorithm and model study, *Journal of Magnetic Resonance*, 1998, **132**, 298-306

94. Antalek, B.; Hornak, J.P.; Windig, W.: Multivariate image analysis of magnetic resonance images with the direct exponential curve resolution algorithm (DECRA) - Part 2: Application to human brain images, *Journal of Magnetic Resonance*, 1998, **132**, 307-315

95. Windig, W.; Antalek, B.; Sorriero, L.J.; Bijlsma, S.; Louwerse, D.J.; Smilde, A.K.: Applications and new developments of the direct exponential curve resolution algorithm (DECRA). Examples of spectra and magnetic resonance images, *Journal of Chemometrics*, 1999, **13**, 95-110

96. Windig, W.; Antalek, B.: Resolving nuclear magnetic resonance data of complex mixtures by three-way methods: - Examples of chemical solutions and the human brain, *Chemometrics and Intelligent Laboratory Systems*, 1999, **46**, 207-219

97. Bijlsma, S.; Louwerse, D.J.; Windig, W.; Smilde, A.K.: Rapid estimation of rate constants using on-line SW-NIR and trilinear models, *Analytical Chimica Acta*, 1998, **376**, 339-355

98. Bijlsma, S.; Louwerse, D.J.; Smilde, A.K.: Estimating rate constants and pure UV-VIS spectra of a two-step reaction using trilinear models, *Journal of Chemometrics*, 1999, **13**, 311-329

99. Bijlsma, S.; Smilde, A.K.: Estimating reaction rate constants from a two-step reaction: a comparison between two-way and three-way methods, *Journal of Chemometrics*, 2000, **14**, 541-560

100. Windig, W.; Antalek, B.; Robbins, M.J.; Zumbulyadis, N.; Heckler, C.E.: Applications of the Direct Exponential Curve Resolution Algorithm (DECRA) to solid stated nuclear magnetic resonance and mid-infrared spectra, *Journal of Chemometrics*, 2000, **14**, 213-227

101. van der Weerd, L.; Vergeldt, F.J.; de Jager, P.A.; Van As, H.: Evaluation of algorithms for analysis of NMR relaxation decay curves, *Magnetic Resonance Imaging*, 2000, **18**, 1151-1157

102. LaTorraca, G.A.; Dunn, K.J.; Webber, P.R.; Carlson, R.M.: Low-field NMR determinations of the properties of heavy oils and water-in-oil emulsions, *Magnetic Resonance Imaging*, 1998, **16**, 659-662

103. Blümich, B.; Blümler, P.; Eidmann, G.; Guthausen, A.; Haken, R.; Schmitz, U.; Saito, K.; Zimmer, G.: The NMR-MOUSE: Construction, excitation, and applications, *Magnetic Resonance Imaging*, 1998, **16**, 479-484

104. Bãlibanu, F.; Hailu, K.; Eymael, R.; Demco, D.E.; Blümich, B.: Nuclear magnetic resonance in inhomogeneous magnetic fields, *Journal of Magnetic Resonance*, 2000, **145**, 246-258

105. Guthausen, A.; Zimmer, G.; Blümler, P.; Blümich, B.: Analysis of polymer materials by surface NMR via the MOUSE, *Journal of Magnetic Resonance*, 1998, **130**, 1-7

106. Somers, A.E.; Bastow, T.J.; Burgar, M.I.; Forsyth, M.; Hill, A.J.: Quantifying rubber degradation using NMR, *Polymer Degradation and Stability*, 2000, **70**, 31-37

107. Zimmer, G.; Guthausen, A.; Schmitz, U.; Saito, K.; Blümich, B.: Weathering investigation of PVC coatings on iron sheets by the NMR MOUSE, *Advanced Materials*, 1997, **9**, 987-990

108. Zimmer, G.; Guthausen, A.; Blümich, B.: Characterization of cross-link density in technical elastomers by the NMR-MOUSE, *Solid State Nuclear Magnetic Resonance*, 1998, **12**, 183-190

109. Haken, R.; Blümich, B.: Anisotropy in tendon investigated in vivo by a portable NMR scanner, the NMR-MOUSE, *Journal of Magnetic Resonance*, 2000, **144**, 195-199

110. Guthausen, G.; Guthausen, A.; Bãlibanu, F.; Eymael, R.; Hailu, K.; Schmitz, U.; Blümich, B.: Soft-matter analysis by the NMR-MOUSE, *Macromolecular Materials and Engineering*, 2000, **276**, 25-37

111. Barfod, N.M.; Krog, N.: Destabilization and fat crystallization of whippable emulsions (toppings) studied by pulsed NMR, *Journal of the American Oil Chemists' Society*, 1987, **64**, 112-119

112. Britton, M.M.; Callaghan, P.T.: Nuclear magnetic resonance visualization of anomalous flow in cone-and-plate rheometry, *Journal of Rheology*, 1997, **41**, 1365-1386

113. Britton, M.M.; Callaghan, P.T.: Two-phase shear band structures at uniform stress, *Physical Review Letters*, 1997, **78**, 4930-4933

114. Guiheneuf, T.M.; Couzens, P.J.; Wille, H.J.; Hall, L.D.: Visualization of liquid triacylglycerol migration in chocolate by magnetic resonance imaging, *Journal of the Science of Food and Agriculture*, 2000, **73**, 265-273

Paper I

Jepsen, S.M.; Pedersen, H.T.; Engelsen, S.B.: Application of chemometrics to low-field ^{1}H NMR relaxation data of intact fish flesh, *Journal of the Science of Food and Agriculture*, 1999, **79**, 1793-1802

Paper II

Bechmann, I.E.; Pedersen, H.T.; Nørgaard, L.; Engelsen, S.B.: *Comparative chemometric analysis of transverse low-field ^{1}H NMR relaxation data*, in "Advances in Magnetic Resonance in Food Science" (Belton, P.S.; Hills, B.P.; Webb, G.A., Eds.), pp. 217-225, The Royal Society of Chemistry, Cambridge, 1999

Paper III

Pedersen, H.T.; Munck, L.; Engelsen, S.B.: Low-field ^{1}H nuclear magnetic resonance and chemometrics combined for simultaneous determination of water, oil, and protein contents in oilseeds, *Journal of the American Oil Chemists' Society*, 2000, **77** (10), 1069-1076

Paper IV

Engelsen, S.B.; Jensen, M.K.; Pedersen, H.T.; Nørgaard, L.; Munck, L.: NMR-baking and multivariate prediction of instrumental texture parameters in bread, *Journal of Cereal Science*, 2001, **33** (1), 59-69

Paper V

Pedersen, H.T.; Bro, R.; Engelsen, S.B.: *SLICING - A novel approach for unique deconvolution of NMR relaxation decays*, in "Magnetic Resonance in Food Science: A view to the Future" (Webb, G.A.; Belton, P.S.; Gill, A.M.; Delgadillo, I., eds.), pp. 202-209, The Royal Society of Chemistry, Cambridge, 2001

Paper VI

Pedersen, H.T.; Berg, H.; Lundby, F.; Engelsen, S.B.: The multivariate advantage in fat determination in meat by bench-top NMR, *Innovative Food Science and Emerging Technologies*, 2001, **2**, 87-94

Paper VII

Pedersen, H.T.; Bro, R.; Engelsen, S.B.: Rapid and unique curve resolution of low-field NMR T_2-components by SLICING and trilinear decomposition, *In prep.*, 2001

Paper VIII

Pedersen, H.T.; Ablett,.S..; Martin, D.R.; Mallett, M.J.D.: Application of the NMR-MOUSE to food emulsions, *In prep.*, 2001