Project Report No. 401

August 2006

Price: £7.50

Managing late N applications to meet wheat protein market requirements using pre-harvest near infrared (NIR) sensing (LK0927)

by

DG Bhandari[1], SJ Millar[1], RM Weightman[2], T Verhoeven[2], JC Richmond[3], PR Shewry[4], DMR Georget[5] and PS Belton[5]

[1]Campden & Chorleywood Food Research Association, Chipping Campden, Glos., GL55 6LD
[2]ADAS UK Ltd, Boxworth, Cambs., CB3 8NN
[3]Bruker Optics Ltd, Coventry, CV4 9GH
[4]Rothamsted Research, Harpenden, Herts., AL5 2JQ
[5]School of Chemical Sciences and Pharmacy, University of East Anglia, Norwich, NR4 7TJ

This is the final report of a forty-eight month project which started in April 2002. The project was sponsored by a Defra LINK project (£214,302, Project No. LK0927) and the HGCA (£171,147, Project No. 2579), plus in-kind contributions from Bruker Optics Ltd (£135,000) and Heygates Ltd (£5,000) making a total of £525,449.

TABLE OF CONTENTS **Page**

1. ABSTRACT

Optimising late nitrogen (N) fertiliser application to achieve breadmaking wheat protein targets remains a problem for growers. Correct use of N fertilisers is imperative for financial and environmental reasons. Given the increasing costs of N fertilisers, there is an urgent need to improve the efficiency of their use. Studies have suggested that the most effective boost to protein comes from using foliar urea at the milky-ripe stage of grain development (Growth Stages 71-75). However, the current basis for late N decisions is inadequate, as only 1 in 4 of the crops benefit directly from extra N in attaining the 13% protein premium. Unfortunately, growers are unable to assess the N status of their crops in order to target fertiliser decisions.

This 4-year project aimed to develop a rapid method for on-farm measurements using new portable FT-NIR (Fourier transform near infrared) technology for optimising N input. In addition, the project sought to increase the scientific understanding of environmental and genetic factors that influence the deposition of proteins that govern end-use quality of wheat.

Wheat was sampled at around GS 70 and at harvest, from trials sown at Boxworth, Terrington, High Mowthorpe, Rosemaund and Essex from 2002 to 2005. The crops received no late N fertiliser and included **nabim** Group 1- 4 varieties, featuring a range of protein contents. The Matrix-I NIR instrument (Bruker Optics) was evaluated for measuring coarsely chopped immature ear and whole plant material. The need for sample drying and, in particular, the effect of microwave drying was investigated. Reference moisture and protein contents (Kjeldahl method) were determined. Protocols were developed for field sampling, as well as for presenting samples for NIR measurements. NIR calibrations were developed and validated for the different forms of wheat material.

Results showed that:

- Using the Matrix-I NIR instrument it was possible to make rapid determinations of protein and moisture contents of fresh (undried), developing ear and whole plant material.
- The most accurate prediction of protein content was generally achieved from the use of ears rather than whole plant material.
- The most accurate prediction of protein (but not moisture content) was generally achieved from the use of dried material.

It was concluded that such NIR assessments could form the basis of an integrated crop monitoring system enabling accurate decision making for the targeted application of late N for boosting grain protein. NIR forecasting of grain protein content at harvest may be improved by using models taking account of the state of grain development and the likely crop yield. This could improve the farmer's ability to meet high quality wheat targets consistently, and improve the marketability and sustainability of the UK arable sector.

2. SUMMARY

In general, farmers face great difficulties in optimising late N inputs to achieve the 13% protein content - dry weight basis (db) target for maximal financial return, as well as minimising diffuse pollution through nitrate leaching. Presently, farmers have no reliable way of rapidly assessing the N status of the crop and specifically identifying those crops which may or may not require extra late N fertiliser reach the 13% (db) protein target necessary for the breadmaking premium. An earlier HGCA-funded study by CCFRA (Bhandari, 2000) demonstrated that it may be possible to predict the protein content and other quality-related parameters from immature grain samples using a laboratory-based near infrared (NIR) instrument. However, despite the highly promising results of the preliminary study, the techniques employed involved freeze-drying, de-husking and grinding of immature grain samples. Clearly, such steps are too slow, highly labour-intensive and impractical.

The aim of this study was to build upon these findings by developing a fast and robust method for measuring immature wheat, in order to help targeting the application of late N fertiliser. In addition, this project sought to provide a better understanding of how genetic and environmental factors affect the accumulation of gluten protein polymers that impact upon the quality of breadmaking wheat. The work was carried out by CCFRA, who were the project co-ordinators. ADAS (Boxworth) managed the growing trials and conducted the field-testing and Bruker Optics developed the prototype portable Fourier transform (FT)-NIR Matrix-I instrument. Rothamsted Research and University of East Anglia (UEA) performed the biophysical studies.

Wheat Trials and Reference Data

This study made use of material from a number of ADAS sites and crops in 2002 and from ADAS-managed N response growing trials, 'N Requirements of Group 1 and 2 wheat varieties', conducted in 2005, for the HGCA (Dampney *et al*, 2006) . In 2003 and 2004, plots were grown specifically for this project, alongside that for the project mentioned above. Further site and experimental details can be obtained from Dampney *et al.*, 2006. These experiments contained a range of varieties, including those from all four **nabim** Wheat Groups, and featured a range of protein contents. The growing wheat received varying levels of early N in the form of ammonium nitrate, but no late N fertiliser. The crop was sampled between the last week in June and the first week in July, at the grain milky-ripe stage (Growth Stage 69-71), from small trial plots (24m x 2m) at sites in Boxworth, Essex, High Mowthorpe, Rosemaund and Terrington. Whole crop samples were cut just above the ground and dispatched to CCFRA the following day. NIR scanning was performed on the fresh whole plant and ear samples. The efficacy of using two different types of microwave oven for sample drying and their impact upon the quality of NIR spectral data obtained were examined. Representative grain samples were obtained from the corresponding trial plots at harvest. The reference measurements were determined using the Kjeldahl method for protein and a gravimetric oven method for moisture.

The ranges of protein (dry matter basis, db) found in the combined data from all sites and seasons were as follows:

Immature whole plant	*3.5 – 12.4 %*
Immature ears	*6.8 – 12.0 %*
Harvest grain	*7.1 – 15.8 %*

The response of the grain protein content to ammonium nitrate fertiliser treatments was particularly affected by site and growing season. Out of a total of 215 harvest grain samples, only 50 achieved a protein content of 13% (db) or higher. The fundamental relationship between the protein in immature wheat material and harvest grain, as measured using a 'wet' chemistry-based method, was influenced by season, growing location, residual soil mineral N and yield. The squared correlation coefficient (R^2) for linear regression was 0.6 for this relationship, and for 82% of the samples the harvest grain protein was correctly predicted as being either below 13% or above 13%. This describes the level of the inherent variation within the reference dataset which was subsequently used for developing NIR calibrations as described in Section 3.3.

Methods were developed for preparing the immature samples prior to scanning by the prototype NIR system at CCFRA. The magnitude of the variation and the sources of variability in using the Matrix-I NIR for protein determination from chopped wheat plant samples were assessed independently at ADAS Boxworth. It was concluded that:

- The major source of variation was due to in-field sampling, and to the time of day they were analysed.
- The variability due to repeated NIR measurements of the same sample was very low.
- Whole plants showed a large proportion of the variation to be due to differences between sub-samples, indicating the importance of proper sub-sampling to gain representative samples.
- Ear samples gave less variable results than measurement of whole plants.
- Chopping plants lowered the variability, particularly the whole plants samples. Fineness of chopping had no significant effect on the mean values estimated by NIR.

NIR Calibrations

NIR spectra were acquired from immature ear and whole plant material using the portable Bruker Matrix-I in 2002, 2003 and 2004 and the laboratory-based Bruker Vector 22/N-I in 2005. Both instruments shared the same optical platform and produced equivalent spectral data. Chopped samples were presented to both instruments in triplicate (fresh) or duplicate (dried). Comparison of the NIR spectra acquired from fresh and dried ear and whole plant material revealed that the major difference between them was dominance of water regions in the spectrum for the fresh samples. Both the immature heads and whole plants had moisture contents in the region of approximately 55-80%. NIR calibrations were developed at CCFRA using reference data for protein and moisture contents for immature and harvest crops. The NIR spectra and the corresponding mean results for the reference data were brought into a single file within the Bruker Opus

software that was used for developing the calibrations for protein and moisture contents using multiple regression analysis. In the first instance, a cross validation approach was taken to assess the calibration performance using the entire dataset. Samples were removed iteratively from a given calibration set, predicted using the remaining samples and then re-introduced to the set. The combination of prediction errors from iterations is called the standard error of cross validation (†SECV). Two additional validation approaches were undertaken. The first of these used samples grown at the Terrington site in all years as the validation set, the results of which were predicted using a calibration derived from the remaining samples for the other sites across all years. The same approach was taken for the second additional validation set in which all the samples (from all sites) from the 2004 harvest year were predicted using calibrations developed for all sites from the remaining harvest years (2002, 2003 and 2005). Comparisons were then made of the performance of the various calibrations developed for determining or predicting the composition of the material assessed (e.g. immature ear protein from spectra derived from immature ears) as well as those used to forecast grain protein content at harvest (using spectra of immature crop material).

It was found that for moisture content, the performance of those calibrations derived using undried material was better than those where microwave oven drying had been used. The results obtained for protein content (db) were marginally better for dried ear and whole plant material than those for the corresponding fresh samples. The calibration performance which sought to forecast mature grain protein content from spectral data collected from immature plant were variable and, generally, not as good as those for developed to determine the actual composition of immature material.

Previously, CCFRA had developed NIR calibrations for predicting protein quality of whole and ground grain and white flour samples as analysed by the Profilblé® size-exclusion high performance liquid chromatography method (SE-HPLC). Details of this study can be found in Project Report 310. The ratio of monomeric gliadin to polymeric glutenin proteins as derived from these analyses is associated with dough strength and breadmaking properties. Such calibrations are considered to be sufficiently accurate to allow them to be used to indicate protein quality (they are currently used for sample assessment in Recommended and National List testing). However, they do not approach the levels of performance that would be expected for determining wheat protein content by NIR. In the present study, calibrations were developed for SE-HPLC analyses that were carried out on harvest grain samples obtained from 2003 and 2004, using spectra from the corresponding fresh and dried immature ear and whole plant samples.

Preliminary assessment of the calibrations developed for the key ratios of SE-HPLC parameters showed that the R^2 values obtained were small (0.03 - 0.45) indicating poor correlations when using a cross validation calibration approach. When examining the results in more detail, however, the SECVs obtained for immature wheat were slightly better than those derived for flour in the previous study. More specifically, those for immature ears appear particularly interesting, and, while the calibrations derived would not be generally applicable at this stage of development, the underlying trend appears worthy of further study.

The main conclusions drawn for the development of the various NIR calibrations are:

- Determination of the protein and moisture contents of fresh (undried, immature) wheat ear and plant material may be performed using NIR.
- The most accurate prediction of protein content was generally achieved from the use of ears rather than whole plant material.
- The most accurate prediction of protein (but not moisture content) was generally achieved from the use of dried material.
- The performance of calibrations seeking to forecast grain protein at maturity was variable under validation conditions.
- Calibrations for protein quality, based on SE-HPLC analyses, showed poorer performance than those for crude protein on a dry matter basis, although evaluation of true performance was restricted by limited sample range.
- The error associated with NIR determination of crude protein (CP) in the immature grain was less than the variability of protein content in the dataset.

†*SECV is the error between the NIR predicted and observed data for the calibration samples calculated by predicting values for a number of subsets of the samples from observations for the remaining samples. More accurate calibrations have lower values relative to the mean and range of the reference data*.

Biochemical Study of Functional Proteins

Two aspects of the protein content of cereal seeds contribute to processing quality, the total amount (protein content) and the composition and properties (protein quality). The behaviour and interactions of the gluten proteins, especially their ability to form high molecular weight glutenin (HMW-G) polymers, are particularly important in determining grain quality. However, the scientific basis underlying the relationship between the events that influence assembly of key polymeric proteins and end-use quality of the wheat crop is not fully understood. While genetic and environmental factors are known to impact upon the composition and structure of the gluten protein during the various stages of grain maturity, the details remain unresolved. Although the main thrust of this project was the provision of a rapid FT-NIR method for predicting wheat protein and quality in harvest grain, it presented a valuable opportunity to exploit the material and data by conducting parallel and complementary biochemical studies on developing wheat proteins.

Work at Rothamsted, in collaboration with UEA and CCFRA, focused on the impact of genotype and nutrition on protein quality. It sought to define these effects and to develop simplified methods for their determination prior to grain harvest. This required the development of new methods of isolating functionally active "gluten protein" fractions from developing grain, in order to relate their properties to the genotype, nutritional status and the end-use quality of the mature grain. Methods for the purification of target proteins were specifically selected in order minimise alteration of their native structure and original properties.

The first method concerned the isolation of protein bodies from the developing wheat kernels of the varieties Cadenza, Riband and Xi19, at about 21 days after flowering. Protein bodies are discrete deposits containing glutenin and other storage proteins. These are found mostly within the aleurone layer and the starchy endosperm of developing grains. During the later stages of grain maturation, it is believed that they form the continuous proteinaceous network embedding starch granules, and which participate in the formation of the gluten matrix in doughs. A method was developed for preparing protein bodies that were suitable for spectroscopic and biomechanical measurements. However, this method was time-consuming, producing variable yields of purified fractions and lower than expected proportions of gliadin proteins.

The second method involved the isolation of a fraction known as "gel protein" which correspond to hydrated glutenin polymers and are enriched in high molecular weight glutenin (HMW-G) subunits. Results from UEA indicated that these preparations were not suitable for spectroscopic and biomechanical measurements due to their high and variable contents of starch.

A third method was developed for isolating gluten protein through the use of a cocktail of digestive enzymes to remove starch and cell wall polysaccharides, without the need for mixing or homogenisation. This gave better and more reproducible yields compared to the other two methods, outlined above. This novel method produced substantially pure fractions which were representative of glutenin polymers that are present at all stages of development, from 14 to 42 days after flowering. Examination of these fractions, and those representing the total protein extracted from developing and mature grain from the field trials at ADAS Boxworth showed that the HMW-G subunits increased with maturity. Although there was evidence of a varietal effect for the relative proportion of HMW-G subunit measured in developing grain samples of the varieties Cadenza and Xi19, no effects of early N fertiliser application were seen.

Spectroscopic Study of Gluten Proteins

Studies were conducted at UEA to investigate the effect of N fertiliser application on wheat quality by spectroscopic examination of the structure of gluten proteins. FT-IR (Fourier transform infrared) and FT-Raman were employed as they are non-invasive techniques which are useful in the detection of changes in the secondary structure in proteins. FT-IR spectroscopy can provide information on protein structural elements such as α- and β-helices, β-turns, β-sheets and random coil. Regions of FT-Raman spectra can be assigned to various aliphatic groups, arabinoxylans, -CH groups of lipids, phospholipids, starch backbone structure and disulphide bonds.

Experiments were performed to characterise such secondary structure features of functional wheat proteins and the impact of factors such as N input and variety during the various stages of development. FT-IR analysis of thin sections of immature grains of Riband and Xi19, treated with a range of N inputs, proved to be inconclusive due to poor reproducibility.

Similar FT-IR results were obtained with protein body preparation, extracted from immature grain samples (representing "unworked" gluten proteins). Better results were obtained when the gel protein preparations of these immature wheat samples were scanned by FT-IR. While Principal Components Analysis (PCA) of the spectral data revealed an effect due to variety, the effect of N fertiliser treatment was not seen. FT-IR analysis was carried out on flour derived from harvest grain samples of Einstein, Hereward, Hereward, Xi19 and Option treated at 2 levels of N fertiliser. The relative proportion of the FT-IR spectral signals assigned to structural components such as β-turns, β-sheets, α-helices and random coils were largely unchanged despite the increase in N fertiliser input.

The low signal to noise ratio was a major feature in the part of the study using FT-Raman spectroscopy. Despite this limitation, it was still possible to assign the various absorbance peaks to specific components. Inspection of the FT-Raman spectra provided additional information relating to protein structure and also information emanating from other components of flour. These included arabinoxylans, phospholipids, ferulic acid and starch. In the case of the variety Option, the analysis of mature flour samples indicated that arabinoxylan and lipid levels decreased with increasing N fertiliser input. This finding was consistent with those reported by other research groups.

Overall Conclusions

This study has demonstrated successfully that the prototype Bruker Matrix-I is a robust FT-NIR instrument that can accurately measure protein and moisture contents in growing wheat. The error associated with NIR determination of protein was relatively low and will give more reliable estimates than those methods that are currently available. This project has successfully shown that:

- It is possible to make useful NIR measurements, rapidly and simply, using fresh, undried samples of wheat.
- The portable Matrix-I can give 'on the spot' readings, giving the possibility for real-time advice on nutrient applications.

The current project has developed protocols for field sampling as well as for measuring samples of the developing crop by NIR with high moisture levels. It has shown that the use of the NIR data alone is insufficient to make accurate forecasts of the grain protein at final harvest, particularly when used across different sites and seasons. This conclusion is entirely consistent with the observation made in the reference dataset. It was seen that the fundamental relationship between the protein in immature wheat and in harvest grain, as measured using the Kjeldahl method, was influenced by seasonal factors, growing location, residual soil mineral N, delayed harvest and crop yield. It is recognised that high yield potential tends to dilute grain protein. Therefore, in order to forecast accurately the final grain protein content, additional prediction of dry matter accumulation is required. It is clear that further improvements in the robustness and general applicability in such a scheme would result from integration of NIR predictions with knowledge of crop development between the time of NIR analysis and harvest. Further work will be necessary to exploit the

positive outcomes from the current project. It is envisaged that by combining the analytical capability of the Matrix-I NIR with a crop modelling approach, it would be possible to predict grain yield and protein content, as a basis for making recommendations on late N applications to breadmaking wheat.

Implications

The milling and baking industry is increasingly delivering a wider range of high quality baked products to meet consumer demand. In order to supply the high quality UK-grown wheats required by the milling and baking industries, there is an increasing need to tailor nutrient management regimes for specific varieties and locations, to meet those needs. Potentially, the measurements made using the Matrix-I could provide a sound basis for controlling key quality attributes through the wheat supply chain. At present, growers have no way of assessing the N status of the growing crops in order to target fertiliser input where benefit can be gained. Successful adoption of a crop modelling approach based on the Matrix NIR system would provide a risk management tools for agronomists and advisors to farmers. This could improve their ability to meet high quality targets consistently, as well as improving the marketability and the sustainability of UK wheat production.

3. TECHNICAL DETAIL

3.1 Description of Reference Dataset

Dhan Bhandari, CCFRA

3.1.1 Introduction

The previous HGCA-funded study by CCFRA (Bhandari, 2000) had demonstrated that NIR can be used for predicting protein content (see Annex III, Figure A1) and several quality-related parameters, including HMW-glutenin content, of harvest grain from analysing immature grain. In this study, spectral data were obtained using the Foss NIRSystems 6500, a laboratory-based instrument. The growing crop samples required a high level of processing for this measurement and so they were freeze-dried, de-husked and milled prior to scanning in the NIR instrument. In addition, only a relatively small data set, limited to three breadmaking varieties grown over two seasons, was used to develop these NIR calibrations. A prototype version of the portable Matrix-I FT-NIR (Bruker Optics) was identified as being a particularly suitable instrument for the provision of a rapid in-field method for assessing the growing crop. The integrating sphere sampling cup assembly of the Matrix-I enables the analysis of coarsely chopped material, without having to grind the sample. The objectives of this part of the study were to produce:

1) Streamlined and more practical protocols for treatment of immature wheat plant material for using the portable NIR instrument for on-farm analyses.
2) Reference data for developing new NIR calibrations (see Section 3.4) using a wider range of varieties, grown at several sites and over several seasons.

3.1.2. Materials and Methods

3.1.2.1 Trial Samples

2002 Samples

A total of 52 samples, including commercial crop, were taken from four sites managed by ADAS in 2002. Crops were selected to cover a range of protein contents. None of the sampled crops received any foliar urea-N treatment. Table 1 summarises the samples according to **nabim** Group classification and the date of sampling of the immature samples. The varieties from the 4 growing sites were:

Boxworth	*Chablis, Charger, Claire, Equinox, Hereward, Malacca, Option, Riband, Soissons, Solstice, Tanker and Xi19*
Rosemaund	*Charger, Claire, Consort, Hereward, Option, Savannah, Shamrock, Tanker and Xi19*
High Mowthorpe	*Claire, Eclipse, Hereward and Xi19*
Terrington	*Chablis, Claire, Consort, Hereward, Malacca, Savannah and Tanker*

The following procedure was undertaken at each of the sites. Representative whole plant samples were cut about 5cm above ground level from individual trial plots (24m x 2m) and from selected areas (~100m^2) of commercial plots. Each sample was about 3kg, consisting of 150-200 tillers, and packed inside sealed plastic bags and stored in a cool environment at 4°C prior to despatch on the day of sampling by overnight courier to CCFRA (Chipping Campden). Harvest grain samples (0.5kg) were collected from each of the corresponding plots.

Table 1. Crop sampling in 2002

Site	**nabim Group**					**Date of sampling of immature plants**
	1	**2**	**3**	**4**		
Boxworth	4	5	2	3		24th June
Terrington	5	1	2	3		26th June
H Mowthorpe	6	0	6	0	+2 organic wheats	1st July
Rosemaund	4	4	3	2		2nd July
(Total)	(19)	(10)	(13)	(8)	(2)	

2003 Samples

A total of 63 samples were taken from trials that were conducted in parallel to the HGCA-funded trials at Boxworth, Terrington and Rosemaund managed by ADAS in 2003 (Dampney *et al*, 2006). Whole crop was cut just above ground level from 24m x 2m plots. Approximately 1kg of each sample of whole crop was packed inside sealed plastic bags and dispatched to CCFRA by courier. Harvest grain samples (0.5kg) were collected from each of the corresponding plots.

Table 2. Crop sampling in 2003

Site	**Variety**	**Ammonium nitrate (kg N/ha)**	**Date of sampling of immature plants**
Boxworth (24 samples)	Hereward Malacca Option Solstice Riband Xi19	0, 100, 200, 300	29th June
Terrington (19 samples)	Hereward Malacca Option Solstice Xi19	0, 80, 150, 250	22nd June
Rosemaund (20 samples)	Hereward Malacca Option Solstice Xi19	0, 100, 200, 300	1st July

2004 Samples

A total of 80 samples were taken from HGCA-funded trials managed by ADAS in 2004 (as in 2003 - (Dampney *et al*, 2006). Whole crop was cut just above ground level from 24m x 2m plots. Approximately 1kg of each sample of whole crop was packed inside sealed plastic bags and dispatched to CCFRA by courier. Harvest grain samples (0.5kg) were collected from each of the corresponding plots.

Table 3. Crop sampling in 2004

Site	**Variety**				**Ammonium nitrate (kg N/ha)**	**Date of sampling of immature plants**
Boxworth (20 samples)	Hereward	Einstein	Option	Xi19	0, 40, 80, 150, 200	29th June
Terrington (20 samples)	Einstein	Malacca	Option	Solstice	0, 40, 100, 150, 200	23rd June
Rosemaund (20 samples)	Hereward	Malacca	Xi19	Solstice	0, 40, 80, 120, 160	30th June
H Mowthorpe (20 samples)	Solstice	Option	Xi19	Malacca	0, 40, 100, 160, 220	6th July

2005 Samples

A set of 24 immature whole crop and harvest samples of Solstice, Mercia, Avalon and Malacca were obtained from an HGCA-funded trial in Essex managed by ADAS in 2005 (Dampney *et al*, 2006). The 4 varieties were treated with ammonium nitrate at rates of 0, 40, 160, 280 and 340kg N/ha. An additional group of 4 immature whole crop samples of Xi19 were taken from adjoining commercial plots and used in the development the NIR calibrations. The immature crop was sampled on 20th June 2005. Harvest grain was not available from the commercial plots.

3.1.2.2 Sample Treatment

Sample handling protocol

Upon receipt of the immature crop, approximately one-third (by weight) of each sample was chopped into 3-4cm length pieces and NIR scanned as the "fresh" whole plant. The ears were taken from the remaining two-thirds of the sample, chopped into 3-4cm length pieces and scanned as the "fresh" ears using a spring-loaded NIR cell. In all cases for the "fresh" material, 3 replicate NIR scans were performed, whereby approximately 50g of sub-sample was packed into the NIR cell, scanned and returned to the bulk sample and mixed. This was repeated twice. Approximately 80g of each sample was stored overnight at 4°C inside sealed bags. The samples were dried using a microwave oven as described below. Outlines of the sample handling protocols employed for the 2002 trial, and those during 2003-2005 trial years are given in Figures 1 and 2, respectively. The effects of overnight storage on FT-NIR measurements were also examined using 2002 trial

year samples. A portion of the immature samples was stored at 4°C and scanned by FT-NIR the following day.

Sample drying treatments

For the 2002 trials, a subset of the immature material was dried using a conventional drying oven (4 days at 40°C and 2 days at 57°C). Another subset was dried using a 750W domestic microwave oven (Matsui 170TC). Approximately 80g samples were treated for 6 minutes at full power. For the 2003 to 2005 trials, immature samples were dried using the 400W Panasonic Inverter range of domestic microwave ovens. Three models with equivalent specifications were used (NN552, NN553 and A750) to enabling high throughput of samples. A beaker of cold water (250ml) was placed in the middle of the turntable as a heat sink. Whole plant material was dried over 30 min and the ears dried over 40 minutes. The heated water in the beaker was replaced with cold water every 5 min during both treatments. The dried samples were scanned using the Matrix-I the following day. Two replicate scans were carried out, whereby the material was removed from the sampling cell, mixed and returned for the second measurement.

This partially dried material was further dried at 50°C over 3 days in a 20-tray Mitchell Batch Tray Dryer for short-term storage. The fully dried material was ground using a Perten Falling Number KT mill (Laboratory Mill 3100). Moisture content was measured according to FTWG Method 08 and protein was determined by the Kjeldahl method according to FTWG Method 09 (CCFRA, 1999). It was originally proposed that protein would be determined by the Dumas method. However, technical issues arose with the designated Leco instrument in 2003, which were subsequently resolved. The use of the Kjeldahl method was retained for the duration of the project for consistency. Figure 1 shows the outlines of the procedures for the handling of the samples taken from trials in 2002 and Figure 2 shows the procedures employed in 2003, 2004 and 2005.

FT-NIR measurements

Measurements were made with a portable Bruker Matrix-I FT-NIR (Figure 3) in trial years 2002, 2003 and 2004. A laboratory-based Bruker Vector 22/N-I FT-NIR instrument (Figure 4) was used for measuring samples taken from the 2005 trials. Both instruments had an integrating sphere assembly for performing measurements on heterogeneous material using a spring-loaded sampling cell containing a quartz glass window (Figure 5). The cup and cell window were cleaned with a brush after each single measurement. In the case of the fresh crop material, the cell window was cleaned after five replicate measurements by wiping it with a paper tissue wetted with acetone, followed by a water-wetted tissue and then a dry tissue.

SE-HPLC Analysis

Protein quality of the 2003 and 2004 harvest grain samples was assessed by size-exclusion high performance liquid chromatography (SE-HPLC). Details of the method can be found in Project Report 310 (Millar, 2003). The SE-HPLC method used in this work is based on that reported by Morel *et al*. (2000) and the data were used for developing NIR calibrations, which are described in Section 3.3.

Figure 1. Sample handling protocol employed for the 2002 trials

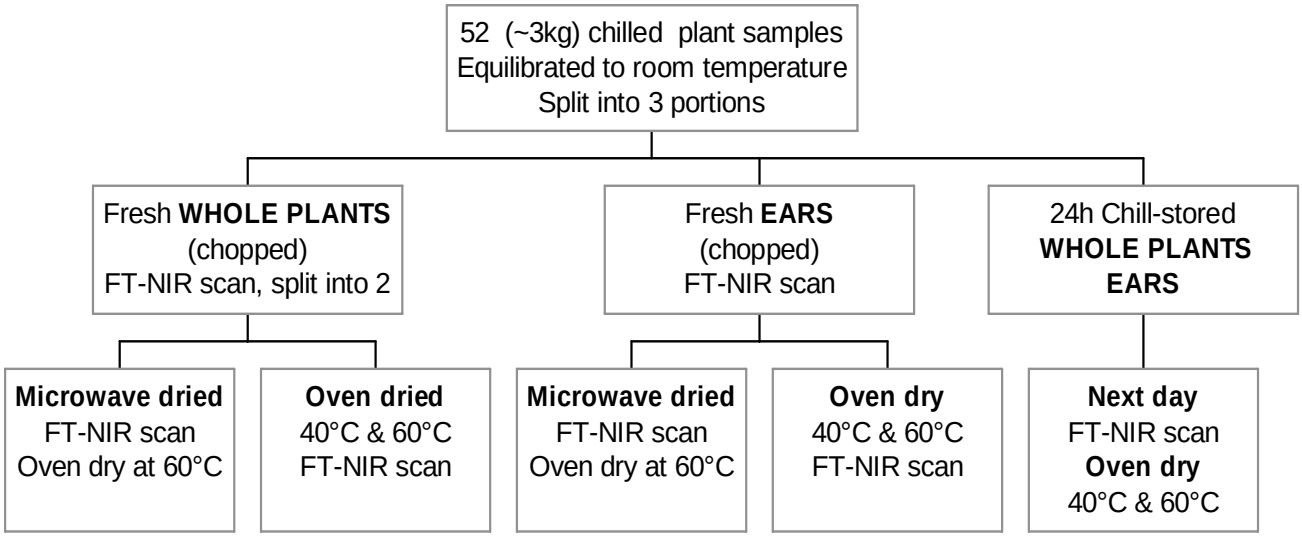


Figure 2. Sample handling protocol employed for the 2003, 2004 and 2005 trials

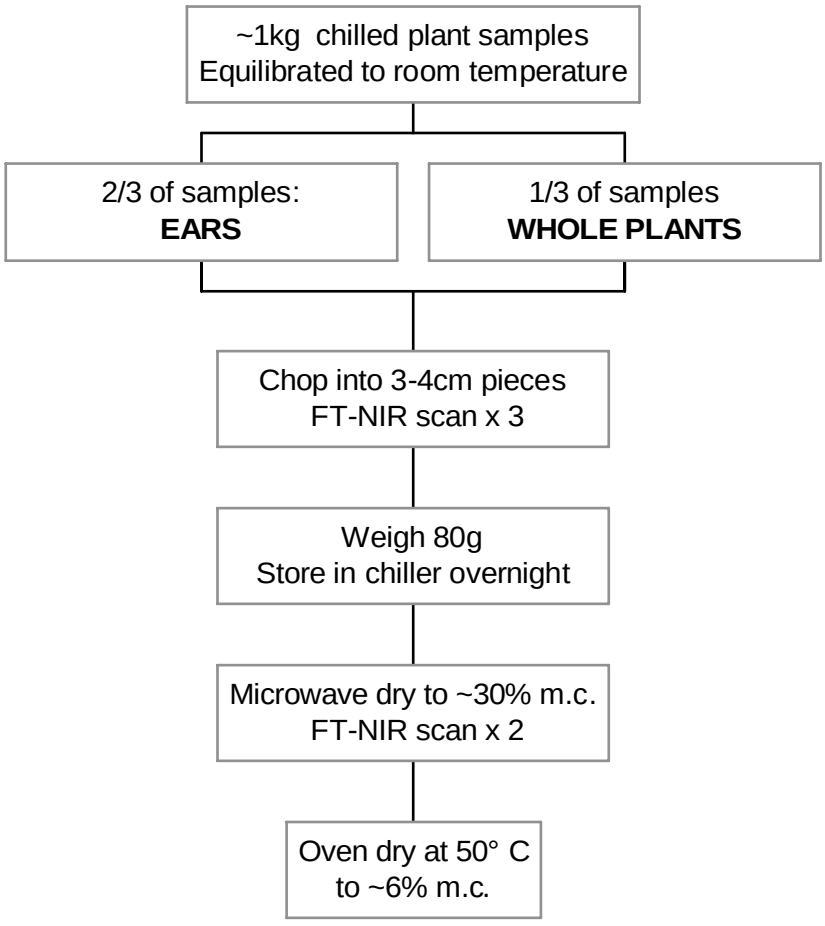


Figure 3. The portable Matrix-I FT-NIR instrument

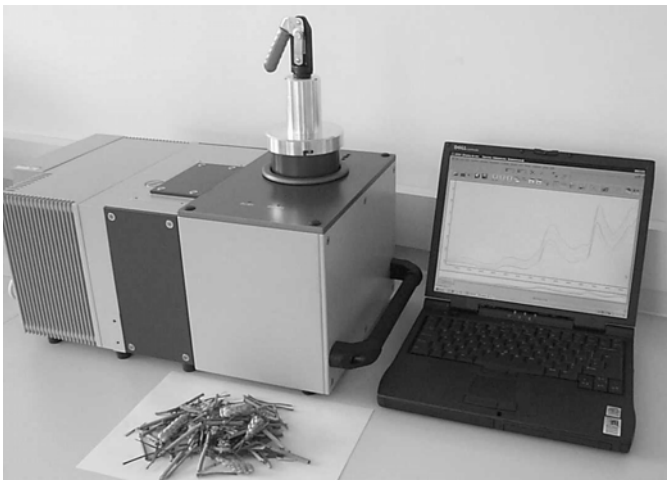
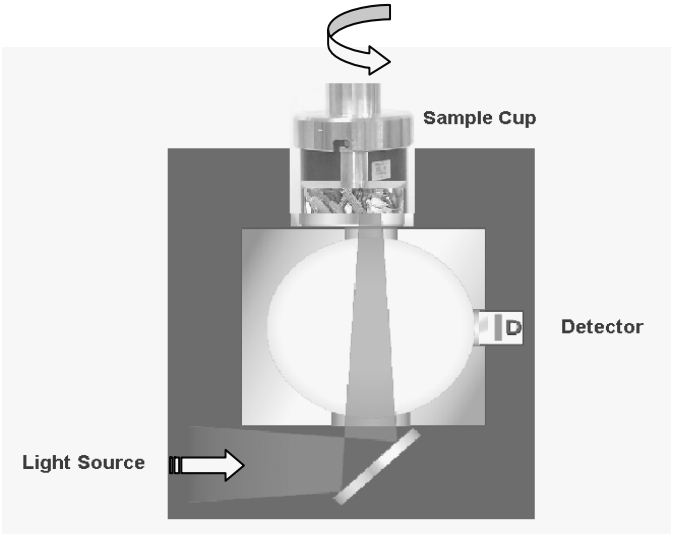


Figure 4. The laboratory-based Vector 22/N-I FT-NIR instrument

Figure 5. The rotating sample cell and integrating sphere detection assembly



3.1.3 Results

Protocols were developed for assessing immature crop samples by NIR, in both "fresh" and dried states. The method deployed for drying "fresh" samples using a standard type of domestic microwave oven (Matsui 170TC) was deemed to be unsatisfactory. While this method was relatively quick and reduced the moisture content to between 15-30%, the Matsui microwave oven generated temperatures as high as 220°C. Thermal imaging of the samples immediately after this microwave treatment revealed localised hotspots; some charring was also observed within the samples. The alternative to this type of oven is a combination-type microwave oven, i.e. with convection heating as well as a magnetron element. The Panasonic 'INVERTER' type of microwave ovens feature low continuous 250W power output and gave a maximum temperature of temperature of 94°C in less than 1% of the sample with an average temperature of 66°C. This method produced moisture contents ranging between 30-40% for the immature wheat material.

3.1.3.1 Relationship between the Protein Content in Immature Crop and in Harvest Grain

The relationship between the protein content (dry matter basis) measured in the immature crop material and that measured in the harvest grain samples was determined for whole plant and ears for the 4 trial years. Tables 4 to 7 contain the squared correlation (r^2) values for trial years 2002, 2003, 2004 and 2005, respectively. The corresponding graphs for these data can be found in Annex I, Figures A1 to A6. A considerable site effect was observed in the 2002 trial year. The highest correlation between crude protein levels in immature crop and in harvest grain were seen at Terrington and High Mowthorpe in 2002 (Table 4 and Annex I, Figure A1). The combined site data for 2002 are shown Annex I, Figure A2.

Table 4. Relationship between mature harvest grain & immature crop protein content – 2002 (Year 1)

Site	Whole plant (r^2)	Ears (r^2)
Boxworth	0.14	0.38
Terrington	0.81	0.79
Rosemaund	0.10	0.08
High Mowthorpe	0.74	0.77
(All sites)	(0.43)	(0.43)

The relationship between measured protein at harvest and in immature crop was more consistent in 2003 than seen in 2002. The combined data for the 3 sites produced a good squared correlation coefficient (0.83) for immature ears and a reasonable value for the whole crop material (0.7). The site effect was greater for the whole plant measurements compared to the ear data. See Table 5 and Annex I, Figures A3 and A4. The relationship between the protein content in harvest grain and in immature crop in 2004 was not as good as that seen in the 2003 trials data (Annex I, Figure A5). The effect of site was evident, whereby Terrington and High Mowthorpe featured the higher (squared) correlation values (Table 6). The corresponding results for the trial site at Essex in 2005 are given in Table 7 and in Annex I, Figure A6. A poor relationship was seen in

this limited set of data from 2005, featuring mostly high protein levels in the harvest grain samples. However, these data are in line with those from the other years.

Table 5. Relationship between mature harvest grain & immature crop protein content – 2003 (Year 2)

Site	Whole plant (r^2)	Ears (r^2)
Boxworth	0.92	0.92
Terrington	0.91	0.81
Rosemaund	0.81	0.96
(All sites)	(0.70)	(0.83)

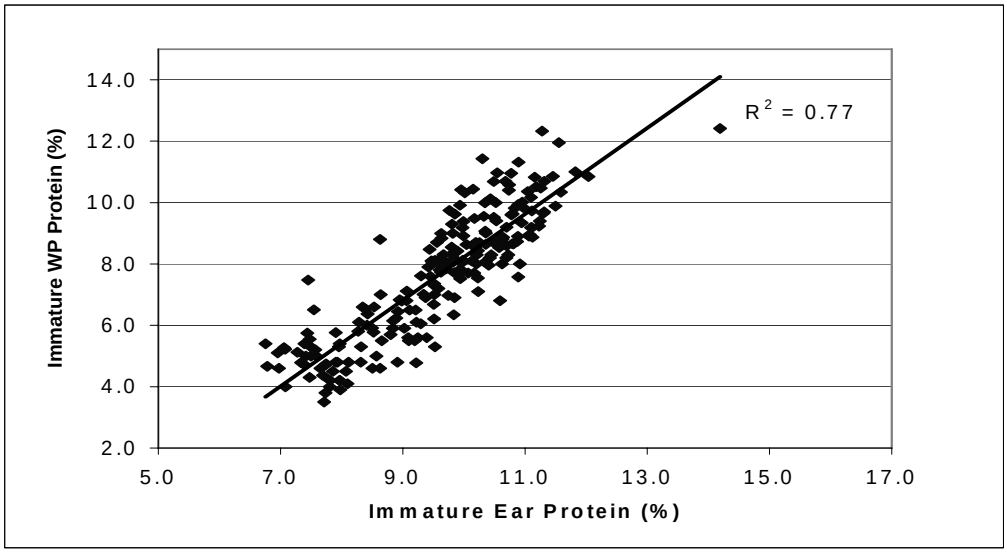
Table 6. Relationship between mature harvest grain & immature crop protein content – 2004 (Year 3)

Site	Whole plant (r^2)	Ears (r^2)
Boxworth	0.67	0.58
Terrington	0.80	0.76
Rosemaund	0.77	0.63
High Mowthorpe	0.90	0.76
(All sites)	(0.41)	(0.53)

Table 7. Relationship between mature harvest grain & immature crop protein content – 2005 (Year 4)

Site	Whole plant (r^2)	Ears (r^2)
Essex	0.29	0.30

Figure 6. Relationship between protein (% db) in immature ears and whole plants (WP) for all sites and all trial years



The combined data for the 4 trial years was plotted to examine the following relationships between the protein content in:

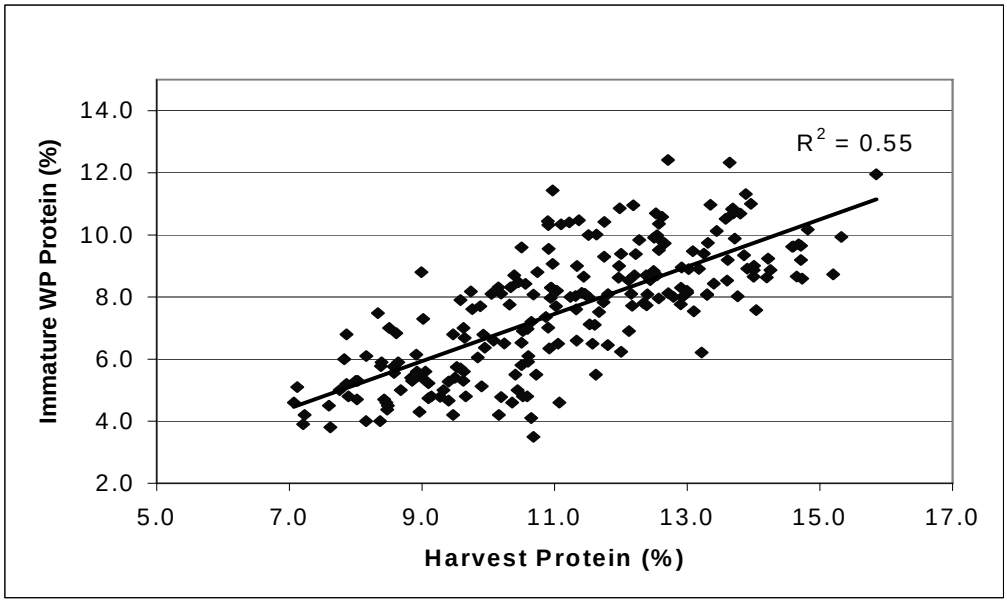
a) Immature ears and whole plants.
b) Immature ears and harvest grain.
c) Immature whole plants and harvest grain.

A goodness of fit (r^2) of 0.77 was found for the relationship between immature ear and whole plant samples (Figure 6) and regression analysis showed that this relationship was not entirely linear (not shown in this report).

Table 8. Relationship between mature harvest grain & immature crop protein content for all sites and all trial years

	Immature Whole Plant + Ears (r^2)	**Immature Whole Plant** (r^2)	**Immature Ears** (r^2)
Harvest grain	0.62	0.55	0.60

Figure 7. Relationship between protein (% db) at harvest and in whole plants (WP) for all 4 trial sites in all trial years



Comparisons of the relationships between the protein content (% db) in harvest grain and in immature ears and in whole plants are shown in Table 8 and in Figures 7 and 8. Detailed analysis of variance of these data revealed that the squared correlation coefficients found for the ears (0.60) and whole plants (0.55) were not statistically different (not shown in this report). Regression analysis of the 215 immature ear and corresponding harvest grain sample data produced the following regression equation:

$$y = -0.485 + 1.22\,x$$

where ***y*** is harvest grain protein (% db) and ***x*** is the immature ear protein (% db). This equation predicts a value of 12.8% for the harvest protein content (equivalent to 13%, db, by the Dumas method) when the immature ear protein content is 10.9%. Figure 9 shows that 50 (23%) of all the harvest grain samples had a protein content (db) higher than 12.8% and that 21 (10%) of these were predicted correctly based on the cut off value of 10.9% for the immature ear protein (db). Similarly, 165 (77%) of samples had protein content lower than 12.8% and 155 (72%) of these were predicted correctly. Therefore, a total of 176 (82%) samples were predicted correctly on this basis.

3.1.3.2 Response to N Applications

The effect of early N fertiliser application, in the form of ammonium nitrate, on the protein content of immature crop (either as ears or whole plant) and harvest grain samples are shown in Annex 1, Figures 9a, 10a and 11a for trials years 2003, 2004 and 2005, respectively. It should be noted that a number of factors affected the crop performance during the 4 trial years. For the 2004 trial year -

a) The harvest was delayed at some sites.
b) The High Mowthorpe site had bad take-all, low yield, shriveled grains and high protein content at harvest.
c) The Boxworth site gave higher protein content than expected at low N input and grain filling was relatively poor compared to the Terrington site.

In the 2005 trials at the Essex site, it was subsequently discovered that the grower had used farmyard manure in the previous season and the soil mineral N was higher than expected. Therefore, the response to applied fertiliser N was poor at this site. Full details may be obtained in Dampney *et al*, 2006.

Figure 8. Relationship between protein (% db) at harvest and in immature ears for all trial sites in all trial years

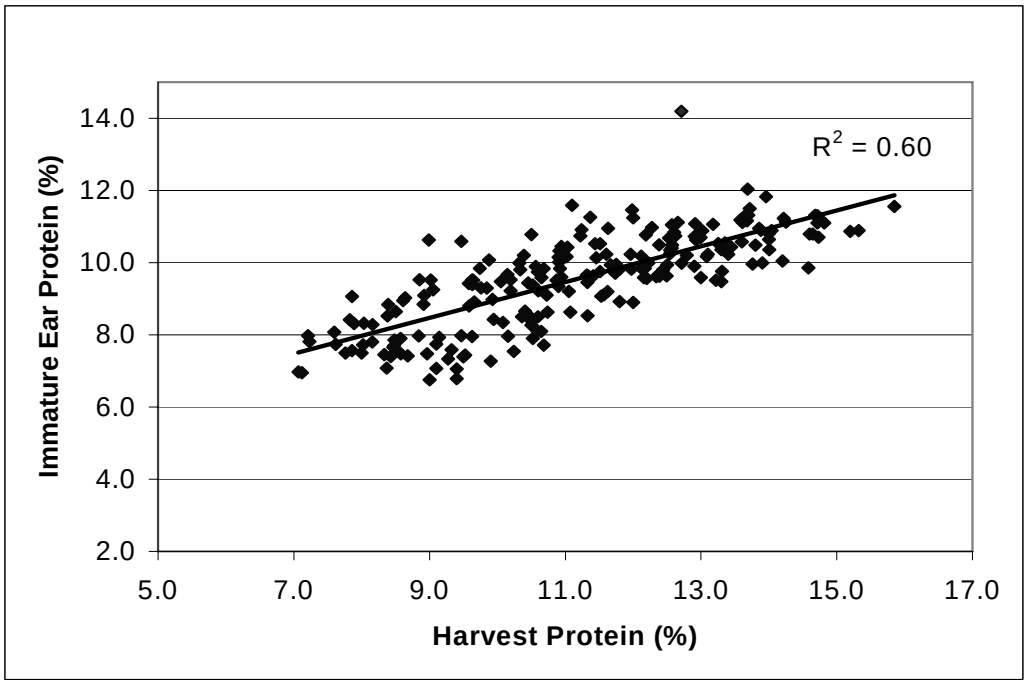
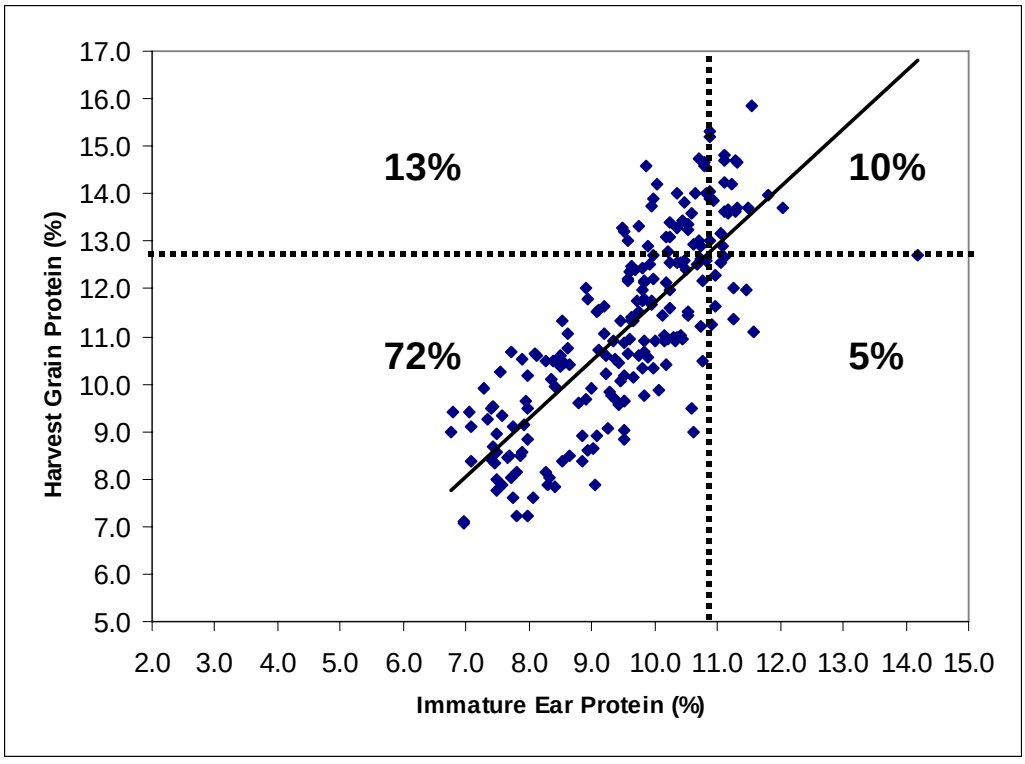


Figure 9. Linear regression prediction of harvest grain protein from immature ear protein (both % db)



3.1.4 Conclusions

- A systematic method was developed for handling and processing immature ear and whole crop that had been sampled at around GS 70.

- An effective protocol was developed for the rapid drying of the immature wheat plant material using a domestic type of microwave oven.

- Effects of site and growing season were found in the relationship between the protein content in harvest grain and immature crop material in the reference data. The protein content of the harvest grain samples ranged from 7.1-15.9%, with only 50 (23%) out of a total 215 samples achieving the premium threshold of 13% (by Dumas).

- For a sample set in which the majority (77%) did not achieve 13% protein, the harvest grain protein was correctly predicted for 82% of the samples.

3.2 Sampling Studies at Boxworth

Richard Weightman and Tamara Verhoeven, ADAS Boxworth, and John Richmond, Bruker Optics

3.2.1 Introduction

Before a system can be used in practice for on-farm measurements, it is necessary to establish the variability in N measurements when using the spring-loaded sampling cell on the Matrix-I NIR. It is important to determine the magnitude of the variation and also to assess the sources of variability that may be associated with each major step of this analytical method. As it was not practical to assess all critical parameters in a single study performed on a single day, the various factors of interest were assessed separately in three different studies as follows:

1. Form of sample study (comparison of ears *vs* whole plant, coarse *vs* finely chopped samples, variability between sub-samples, variability due to repeated measures).
2. Cleaning study (importance of cleaning of cell window, when chopped ears are sampled).
3. Packing study (comparison of cell full/half-full, and for the full cell, packed *vs* unpacked).

This section discusses the results of these studies carried out on a commercial crop of wheat (variety Robigus) at ADAS Boxworth.

3.2.2 Materials and Methods

Samples were collected on 27/6/05 (study 1) and 29/6/05 (studies 2 & 3). Analysis took place on the day of sampling, except in the case of study 3, where 12 of the 48 samples were processed and analysed for protein (db and expressed as crude protein) and moisture by nir on 30/6/05, using calibrations developed by ccfra from data acquired between 2002 and 2004.

3.2.2.1 Study 1: Form of Sample

Structure:

- 8 samples from within the field
- 2 forms of plant materials (whole *vs* ears)
- 2 forms of sample (whole *vs* chopped)
- 2 subsamples of each (replicated on the plant material)
- 2 repeat measures of each sub-sample (replication of the measurement within the cell which must be purely due to analytical variation)

All combinations thereof:

8 x 2 x 2 x 2 x 2 = 128

ANOVA was carried out on the full date set (n=128) and also separately for the whole plant and ears (n=64 in each case). In the case of the 8 plant samples, the data were analysed as 4 sets [effectively blocks] based on time of day measured x 2 samples at each.

3.2.2.2 Study 2: Cleaning the Cell

Five chopped ear samples were analysed using the Matrix-I, after beginning with a clean cell, with the cell getting steadily dirtier. The cell window was then cleaned again after the fifth sample.
Structure:

- 6 samples from within field
- 2 forms (whole *vs* chopped ears)
- 5 measurements (measure after the cell has been cleaned; measure after one sample, uncleaned..... measure after five samples, uncleaned)

6 x 2 x 5 = 60 samples

For the purposes of the ANOVA, the data from the 6 field samples was analysed as 3 blocks (based on time of day measured) x 2 samples each.

3.2.2.3 Study 3: Packing the Cell

Structure:

- 8 samples within field
- 3 packing regimes - 'full cell, unpacked', 'full cell, packed' and 'half-full' cell, packed' [here 'packed' means tension applied to the sample via the piston mechanism in the cell]

2 measurements of each

8 x 3 x 2 = 48

For the purposes of the ANOVA, the data from the 8 field samples was analysed as 4 blocks (based on time of day measured) x 2 samples each.

Analysis method

In each case, ANOVA was used to partition the variance between the different factors. In certain examples, the means and standard deviations are tabulated where these illustrate differences in variability between different factors.

3.2.3 Results

3.2.3.1 Study 1 Results

When all 128 samples were analysed together, it was clear that the largest proportion of the variance was attributable to the difference between ears and whole plant, where the means differed markedly (crude protein 5.87% *vs* 8.17% respectively). Results of this 'full' ANOVA are presented in Table 9.

Table 9. ANOVA for form of sampling study, with all samples (whole plants and ears) analysed within a single structure (n=128)

Stratum	Description	DF	SS	MS	Ratio*
1	Block (Time of day)	3	6	1.85	31
2	Samples (within blocks)	4	6	1.46	24
3	1.2.Form of sample	8	182	22.74	378
4	1.2.3.Fineness of chopping	16	25	1.57	26
5	1.2.3.4.Subsamples	32	114	3.54	59
6	1.2.3.4.5.Repeated measures on Matrix-I	64	4	0.06	1
(Total)		(127)	(336)		

* Ratio of MS for stratum of interest, to MS of lowest stratum (=1)

Conclusions Table 9

- Variability due to repeated measures in the machine was very low.
- Variability associated with strata 1, 2 and 4 were similar in magnitude.
- Greatest variability was associated with differences between whole plant and ears (stratum 3).
- The next important source of variability was that between sub-samples (stratum 5). Later analysis showed that this was mainly due to variability between whole plant sub-samples, rather than ears.

The results suggest that ears give the least variable results, perhaps confirming observations earlier in the project. This can be illustrated by consideration of means and standard deviations (Table 10).

Table 10. Mean crude protein (%) and standard deviations of ears and whole plants, either coarsely or finely chopped.

Sample	Mean	SD	n	SD as % of mean
Ears (average of coarse & fine)	5.87	0.480	32	8.2
Ears coarse	5.91	0.516	16	8.7
Ears fine	5.83	0.446	16	7.7
Whole plant (average of coarse & fine)	8.17	1.555	32	19.0
WP coarse	8.45	1.791	16	21.2
WP fine	7.89	1.244	16	15.6

ANOVAs for the separate plant factions are shown in Table 11 (ears) and Table 12 (whole plant). In general, the variability due to repeated measures or scans of the same sample by the Matrix-I NIR was very low, whichever fraction was analysed. Variability due to blocks and plants always proved major sources of variation. However, the sources of variability differed between the two plant fractions.

For ears, the greatest source of variation was variability between plants. In contrast, ANOVA of whole plant showed most variability between different sub-samples (Table 12).

Fineness of chopping was of relatively minor importance in the case of ears, but was a more important source of variation with whole plant samples.

Table 11. Results of ANOVA for form of sampling study, ears only

Stratum	Description	DF	SS	MS	Ratio*
1	Block (time of day)	3	1.3	0.44	7
2	Samples (within blocks)	4	5.5	1.37	23
3	1.2. Fineness of chopping	8	1.1	0.14	2
4	1.2.3. Sub-samples	16	4.5	0.28	5
5	1.2.3.4.Repeated measures	32	2.1	0.06	1
(Total)		(63)	(14.5)		

Ratio of MS for stratum of interest, to MS of lowest stratum (=1)

Conclusions Table 11

- Variability due to repeated measures of crude protein (%) in the instrument was very low.
- Variability associated with strata 1, 3 & 4 were similar in magnitude.
- Greatest variability was associated with differences between plants (stratum 2).
- Variability between sub-samples (stratum 4) was much less important for ears than it was for whole plants (see next table).

Table 12. Results of ANOVA for form of sampling study, whole plant only

Stratum	Description	DF	SS	MS	Ratio*
1	Block (Time of day)	3	11	3.78	63
2	Samples (within blocks)	4	6	1.46	24
3	1.2. Fineness of chopping	8	24	3.01	50
4	1.2.3. Sub-samples	16	109	6.81	114
5	1.2.3.4.Repeated measures	32	2	0.06	1
(Total)		(63)	(152)		

* Ratio of MS for stratum of interest, to MS of lowest stratum (=1)

Conclusions Table 12

- Variability due to repeated measures of crude protein in the machine was very low.
- Greatest variability was associated with differences between sub-samples (stratum 4) for whole plants.
- Variability between samples (stratum 2) was of similar importance as with ears, but the block (time of day; stratum 1) effect was greater for the whole plant samples.
- Variability associated with strata 1, 3 & 4 for whole plants was greater than that found with ears.

3.2.3.2 Study 2 Results

The cleaning study was carried out using ears only, and aimed to characterise the significance of the presence of contamination from the cell contents of immature grains (which are left smeared on the cell window when ear samples are chopped).

The results suggested that the order of sample processing after cleaning the cell window only made a small contribution to the overall variability (Table 13). Study of the variance ratio (not presented here) would suggest that analysed as a treatment effect, the order of cleaning was not significant. However, it should be noted that with the variety used in the present study, there was relatively little sap expressed when samples were chopped. If samples were taken at an earlier development stage, this might be a more important factor, and should then be revisited.

Table 13. Results of ANOVA for cleaning study

Stratum	*Description*	DF	SS	MS	Ratio*
1	Block (Time of day)	2	4.5	2.25	8
2	Samples (within blocks)	3	1.6	0.54	2
3	1.2.Form of ears	6	0.9	0.14	0.5
4	1.2.3.Order of cleaning	48	13.1	0.27	1
(Total)		(59)	(20.1)		

* Ratio of MS for stratum of interest, to MS of lowest stratum (=1)

Conclusions Table 13

- Variability due to order of cleaning cell was very low.
- Variability due to form of ears (i.e. chopped *vs* whole ear) was very low.
- The major source of variability was the block effect (time of day).

3.2.3.3 Study 3 Results

The packing study was carried out with whole plants only. The ANOVA is shown in Table 14. Form of packing made only a small contribution to the overall variability, whereas the block effect was relatively large. In part this could be due to the fact that block 4 had to be stored in a cold store overnight and analysed the following day. Despite the lack of an overall effect of packing, the SD appeared to be lowest with the full cell, unpacked (Table 15), which is possibly a reflection of the fact that the light within the cell had the longest pathlength with this form of packing. Therefore, better averaging of the light scattered within the chamber was achieved with the full, unpacked cell. It should be noted that all the NIR measurements at CCFRA were performed with approximately half-full cell, packed. This procedure yielded a marginally higher SD, but similar mean protein (%) to that obtained using full cell, unpacked.

Table 14. Results of ANOVA for packing study

Stratum	Description	DF	SS	MS	Ratio*
1	Block (Time of day)	3	4.6	1.52	4
2	Samples (within blocks)	4	1.7	0.41	1
3	1.2.Packing	16	5.3	0.33	1
4	1.2.3.Repeated measures	24	9.6	0.39	1
(Total)		(47)	(21.1)		

* Ratio of MS for stratum of interest, to MS of lowest stratum (=1)

Conclusions Table 14

- Variability due to block (time effect) was the major source of variation.
- Blocks, and packing treatments contributed similar levels of variation.

Table 15. Mean crude protein (%) and standard deviations of whole plants with three different packing regimes

Sample	Mean	SD (n=16)	SD as % of mean
Full cell unpacked	4.11	0.510	12.4
Full cell packed	4.10	0.729	17.7
Half full cell	3.90	0.765	19.6

3.2.4 Conclusions

- In all cases, the major source of variation was between samples from the field, and the time of day they were analysed. In contrast, the variability due to repeated measurements by the Matrix-I NIR of the same sample was very low.

- Analysis of whole plants also showed a large proportion of the variation to be due to variation between sub-samples (i.e. within a sample taken from the field) indicating the importance of proper sub-sampling to gain representative samples.

- The assessment of ear samples appears to give less variable results than measurement of whole plants. However, the decision regarding whether to use ear samples alone depends on the value of ear-N measurements from a physiological point of view. This has not been addressed in the study reported within this section.

- Chopping plant fractions tended to reduce variability, and this was particularly marked with the whole plant samples, but was less important with ears. However, fineness of chopping had no significant effect on the mean values estimated by the Matrix-I.
- Cleanliness of the cell had no effect on mean crude protein values, and only made a small contribution to the overall variability measured. It is concluded therefore that it is not necessary to specify in a protocol that the cell window should be cleaned between every measurement.

- Packing had only a small effect on the mean values for crude protein. However, these and other results suggest that to minimise the variability, the cell should be filled as fully possible with plant material.

3.3 NIR Calibrations

Sam Millar, CCFRA and John Richmond, Bruker Optics

3.3.1 Introduction

3.3.1.1 Background to the use of NIR spectroscopy

The term spectroscopy covers those analytical techniques which rely on the interaction between radiation and matter. The near infrared is that portion of the electromagnetic spectrum which covers wavenumbers from approximately 14285 to 4000cm^{-1}. This corresponds to radiation having wavelengths from 700nm to 2500nm. For reference, the light that human beings can see covers the range 400nm to 700nm. The near infrared region includes radiation which is felt as heat to humans even though our eyes are not able to see it. Molecules having strong dipoles absorb near infrared energy as a function of the different ways in which the molecules vibrate. Typical molecular groups include C-H, N-H and O-H and vibrational modes are those related to stretching and bending. As a result, an NIR spectrum of a given material contains information relating to lipid, protein and moisture primarily although any materials containing molecules having strong dipoles will be likely to make a contribution to NIR spectra.

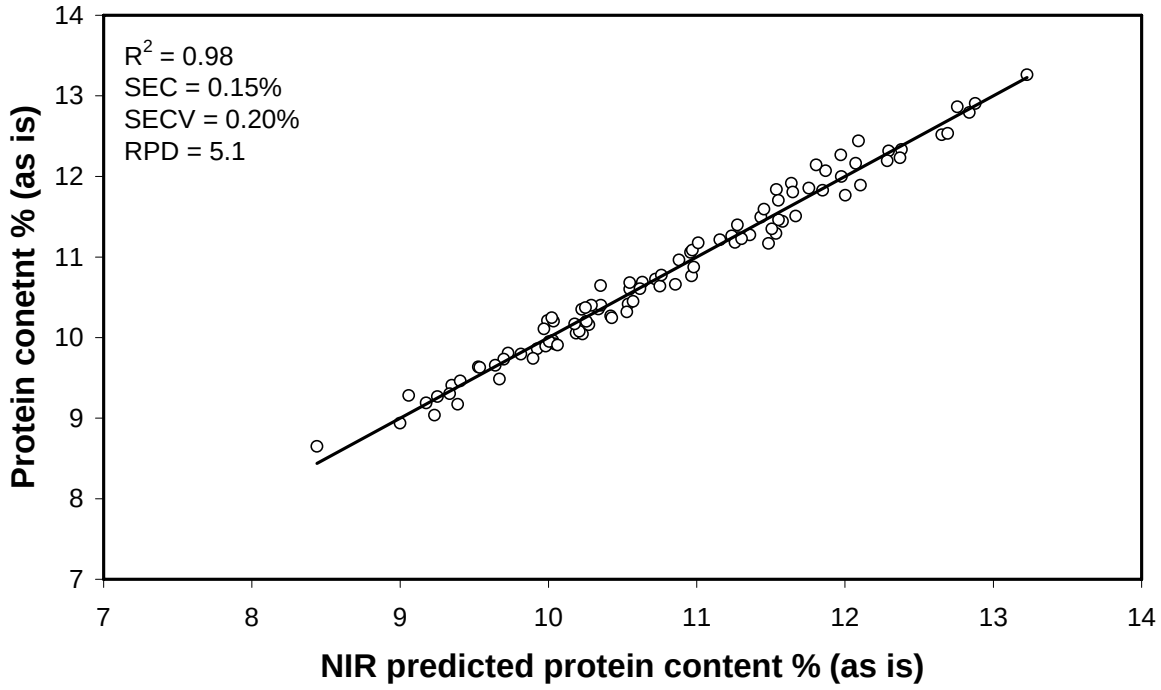
NIR spectra are generated using spectrometers which compare the radiation reflected by or transmitted through a sample with that reflected by or transmitted through a reference. The resulting differences are plotted across the relevant wavenumber range and take the form of a characteristic pattern, the shape of which is indicative of the composition of the material which has been assessed. There are a number of ways in which this information may be collected for processing including interference filters, diffraction gratings and interferometers. The last of these forms the basis of the instrumentation used within the work reported. Spectra are acquired as a summation of a series of sine waves from which the underlying shape of the spectrum may be derived using Fourier transform (FT) algorithms. Collecting spectra in such way can allow for more robust instrumentation (as in this case) but also generates lower levels of noise than is typical for grating systems.

In addition to the absorbance of radiation by specific molecular groups, the physical characteristics of a given sample also affect the spectrum by scattering radiation. As a result, the apparent absorbance at any given wavenumber cannot be directly related to the composition of a given component in a system. Rather, a multi-component calibration needs to be derived with a range of wavenumbers being included to account for physical differences between spectra as well as the variation in the chemical components of interest. The basis of NIR calibration, therefore, is the correlation of the spectral information with a reference measurement for the component of interest such as protein content. In modern calibration approaches, the entire spectrum for each sample is used during calibration development through multivariate data reduction processes such as Principal Components Analysis (PCA) and partial least squares (PLS).

3.3.1.2 Measures of NIR Calibration Performance

The success of any calibration process may be determined by evaluating the closeness of the fit between the values actually obtained. Figure 10 gives an example of this process for an NIR calibration for protein in mature wheat.

Figure 10. Example NIR calibration for protein content in mature wheat, abstracted from HGCA Report No. 310 (Millar, 2003).



There are a number of statistical measures of this performance the relevance of which vary depending on the requirements of the operator. The most universally understood is the squared correlation coefficient (r^2 or R^2 for multiple regressions as typically used for NIR data). This is the proportion of the overall variation in the relationship between the correlated values which is explained by the derived regression equation or model. Values approaching 1 indicate a very strong correlation to have been generated. This has the advantage that it is widely understood and is dimensionless such that comparisons may be made between different parameters which may have very different values. The disadvantage of the approach is that R^2 is dependent on the range of values included. Circumstances may occur, therefore, where two groups of data having strongly differing values are included in a regression model for which only one group is of practical relevance. The inclusion of the second group serves to artificially increase the range and thus the R^2, even though the underlying performance for either group on its own may not be particularly good.

It is generally more useful, therefore, to compare calibrations on the basis of the error of determination associated with them. There are a number of recorded measures of this, each of which is broadly based on the square root of the mean of the squared difference between each pair of compared values.

When generating an NIR calibration, the standard error of calibration (SEC) is derived as part of the calibration process. This serves as a useful indicator of the performance of the model but as the SEC is often used as the response that is deliberately minimised during the calibration process, it leads to an optimistic assessment of calibration performance. A better test of future performance (which is the most important aspect of NIR calibration performance) is the standard error of prediction (SEP). This is calculated using the same basic method as the SEC but is calculated for pairs of values generated for a set of samples which is independent of the calibration set. This is the 'gold standard' for testing NIR calibration performance but is recognised to compromise the calibration process to an extent as any sample in the validation set is not being used to generate a better calibration, simply to test the existing calibration. An alternative method which has become more widely used with the increased use of multivariate calibration approaches is cross validation where within a given calibration set, samples are iteratively removed from the set, predicted using the remaining samples and then re-introduced to the set. This process is repeated for a number of groups of samples until all samples have been removed and replaced. The combination of the prediction errors from iteration is called the standard error of cross validation (SECV). This is generally taken as a more robust indicator of likely future performance than the SEC even if, for many users, SEP remains the ultimate test of an NIR calibration.

Each of these error parameters are expressed in terms which may be equated directly to the magnitude of the values of the parameters predicted. This makes for easy comparison between calibrations for a given component but makes comparison of calibrations for parameters having very different values more difficult. In addition to R^2, therefore, another dimensionless parameter (the ratio of performance to deviation or RPD) may be derived relating the magnitude of the errors in the variability or range of the values recorded to the calibration's performance. Values of above 1 indicate that the standard deviation of the dataset used for calibration development is greater than the error of the model. In practice, this should always be attained as a value of 1 means that the prediction is essentially random within the variation of the population. The larger the value for RPD, therefore, the better is the calibration.

The magnitude of the errors achieved during calibration is assumed to be normally distributed. Thus, differences in the value obtained from that predicted by NIR may be evaluated on the basis of whether they fit the normal distribution of the population to which they belong. This is also the reason why the use of an NIR calibration, even one with relatively poor performance, would be expected to be better than simply estimating the value of a given sample based on the underlying distribution of the population.

Figure 11 shows an idealised situation where it is assumed that the values for wheat protein in a population are normally distributed. The majority of samples are close to the mean value with progressively fewer and fewer samples occurring as one moves to the extremes of the population. For a given field, the values for wheat protein will fall somewhere within the population but without other information, a figure for protein selected at random is unlikely to be the correct one. By estimating a number of values based on the

population, the mean of all the estimates will tend towards the mean of the overall population. The strength of using an NIR approach in such a case is twofold. Firstly, the NIR calibration error should be less than the population standard deviation (i.e. the RPD will be greater than 1) as shown in Figure 12. In addition, however, the mean of a number of determinations will tend towards the true mean of the sample assessed and not the mean of the overall population (Figure 13).

Figure 11. Example of a normal distribution having a standard deviation of 1

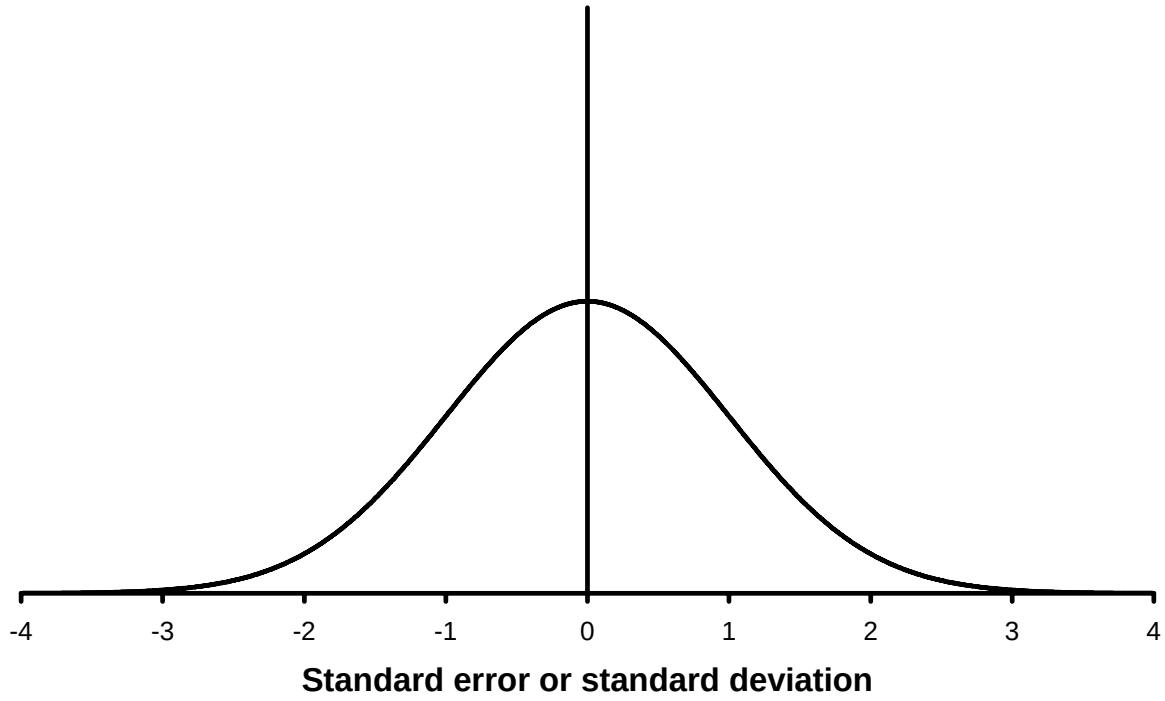


Figure 12. Example of an NIR calibration having a standard error of 0.5

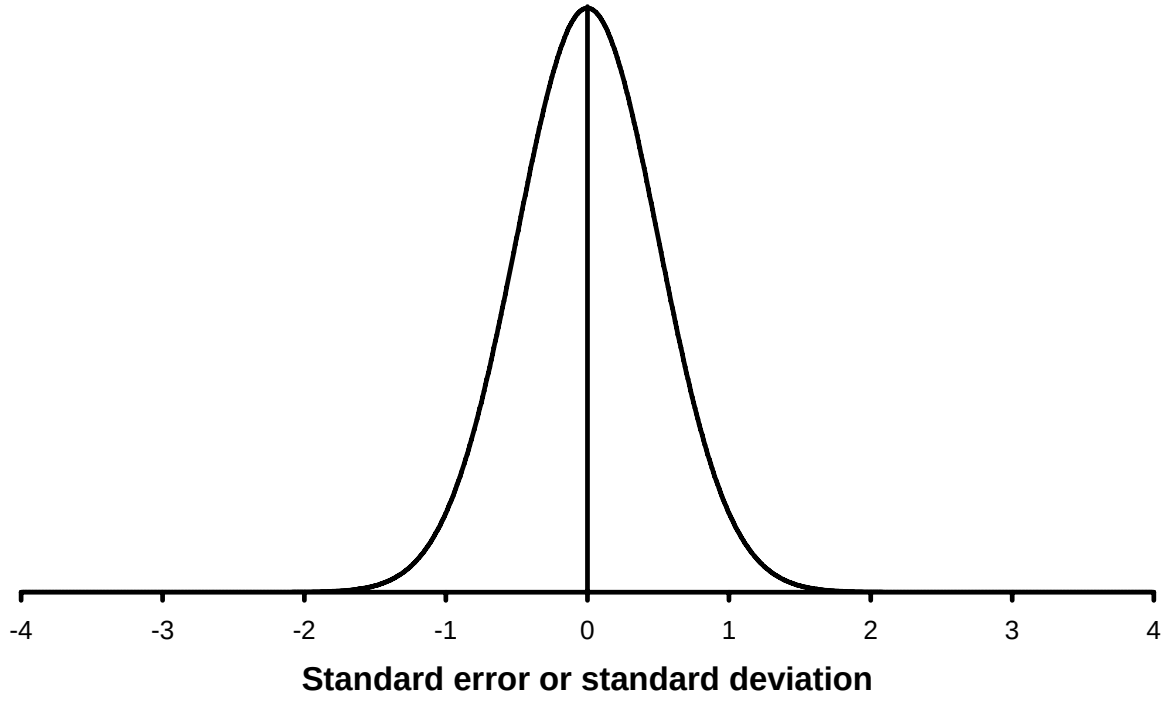
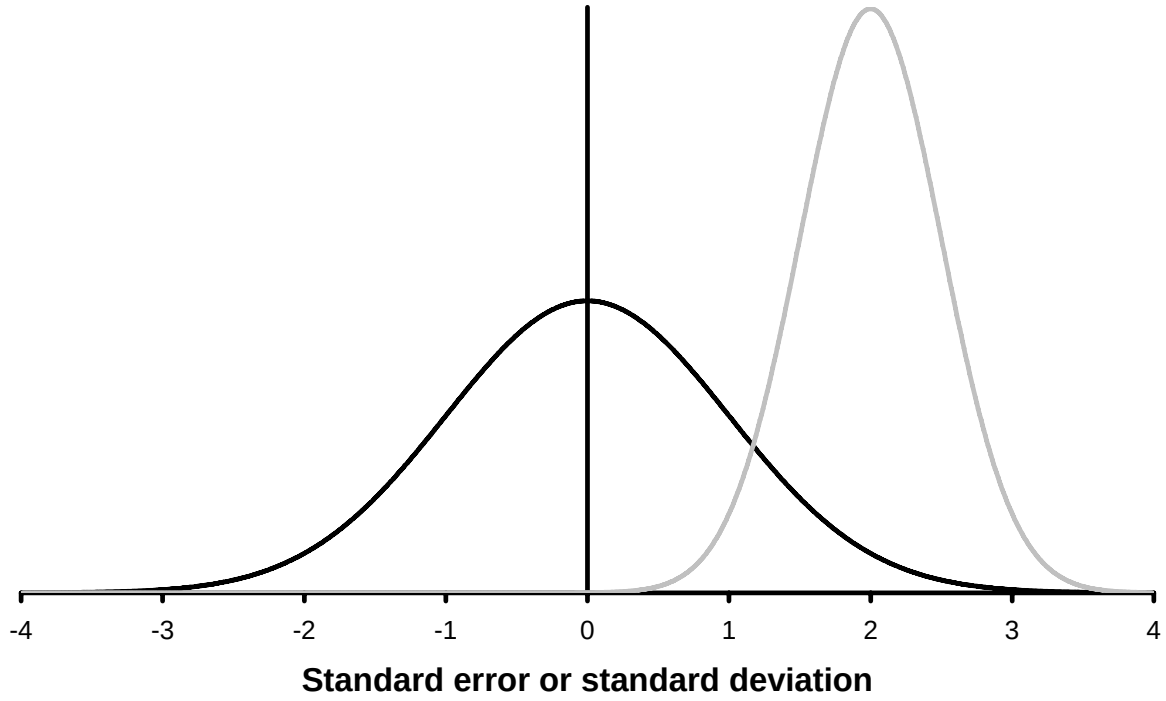


Figure 13. Comparison of the distribution of errors for a population having a standard deviation of 1 and that of an NIR calibration having a standard error of 0.5. The means of the two distributions differ.



Assuming, therefore, that NIR may be used to determine protein content, the values thus obtained should be expected to provide a more reliable basis for N decision-making than estimates based on generalized figures for wheat protein.

3.3.2 Materials and Methods

3.3.2.1 NIR Calibration Development

NIR spectra and corresponding reference data were brought into a single file within the Bruker Opus software for calibration development. For the moisture and protein contents of the immature material, the reference data were those which were directly related to the material for which NIR spectra had been collected and constituted determinations performed on sub-samples of the same material. For forecasted protein and SE-HPLC ratios, the reference data used for each spectrum were those derived from mature grain collected from the same plot as the immature wheat used to generate the spectra. In each case, all the replicate spectra (3 and 2 for fresh and dried material respectively) were included in the calibration files and the common reference data for each sample was entered for each individual replicate. Thus for fresh material, each set of 3 spectra per sample had differing spectra but the same reference values for each spectrum. Given the differences between replicate spectra, therefore, the NIR predicted values would not be expected to be the same for each replicate of a given sample even though the reference data were included as identical values.

Prior to each calibration development process, the entire set for each parameter was used to generate calibrations using cross validation with a group size of 100. Graphs of NIR predicted versus reference data were generated in each case and inspected for outliers. Those samples having large residuals (where the relationship between the predicted and reference values was not typical of the general population) were removed and the calibration re-developed. When only samples which fell within the statistical distribution of the entire calibration set remained, final calibration development was undertaken.

Calibrations using all the samples remaining after removal of outliers were developed by cross validation using a group size of 50. The optimum calibration was determined in each case as that having the lowest SECV following calibrations being developed for a fixed number of operating parameters using the Optimise function in the Opus software. In each case, the mathematical pre-treatments which were evaluated were: unprocessed spectra, vector normalisation and vector normalisation followed by first derivative. A fixed series of wavenumber ranges corresponding to the Bruker preset NIR Optimisation were assessed for each of these pre-treatments.

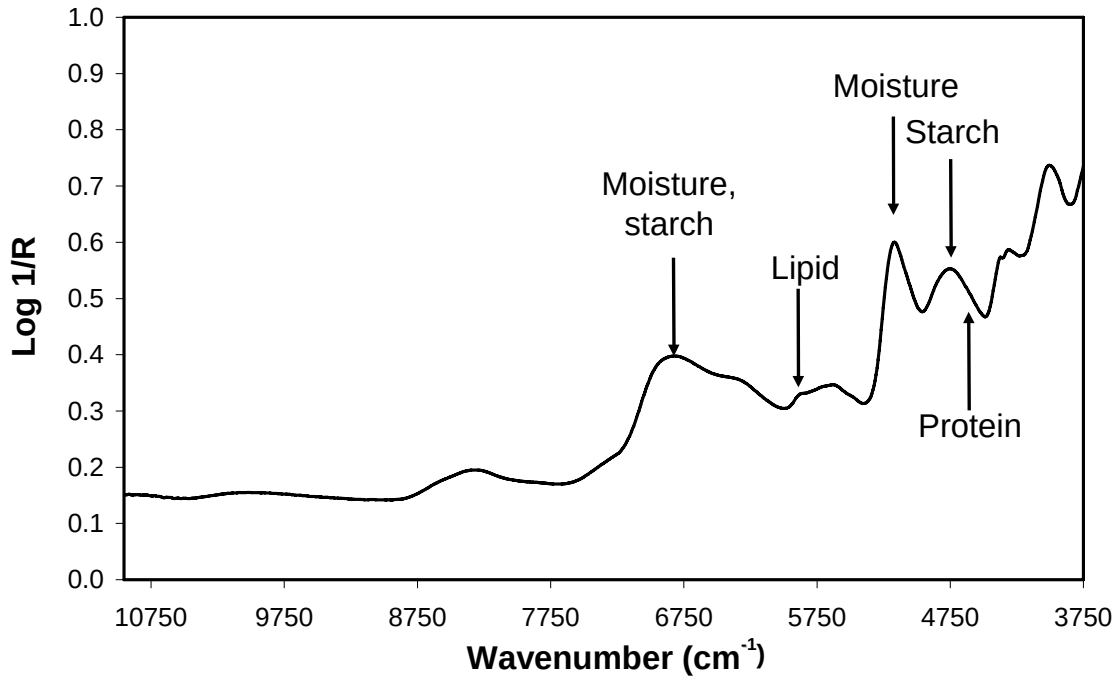
When developing calibrations using a validation set, the same Optimise approach was undertaken. In this case, however, the calibration process did not use cross validation to obtain the best calibration solution. In this case, the standard error of prediction for the validation set was used as the determinant of optimum calibration method.

3.3.3 Results and Discussion

3.3.3.1 NIR Spectra

An example spectrum for ground mature wheat grain is given in Figure 14 with a number of the most important absorbance bands highlighted. It may be seen that NIR spectra tend to have broad and relatively poorly-defined features compared with mid infrared spectra (see Section 3.5). This is due to the information in the shorter wavelength region being the overtone and combination bands from the primary absorbance bands which are found at lower wavenumbers. This means that NIR spectra inherently contain less-detailed information than the mid infrared region, which may be used to determine detail such as the conformation of specific bonds. The usefulness of NIR for agricultural materials, however, is due to the less intense absorbance experienced in this region which allows bulky materials containing higher levels of moisture to be assessed with minimal transformation. Such is the response to water in the mid infrared region that the useful penetration of the incident radiation is to the depth of a few microns making it impossible to derive a meaningful spectrum from the material used in this study for near infrared measurement. Nevertheless, the ability to determine information related to bond conformation means that mid infrared spectroscopy has been applied successfully within the work reported (Section 3.5) using the specialised technique of attenuated total reflectance (ATR). This allows spectra to be collected on the basis of a limited interaction of radiation with a small amount of material.

Figure 14. Example FT-NIR spectrum of ground mature wheat



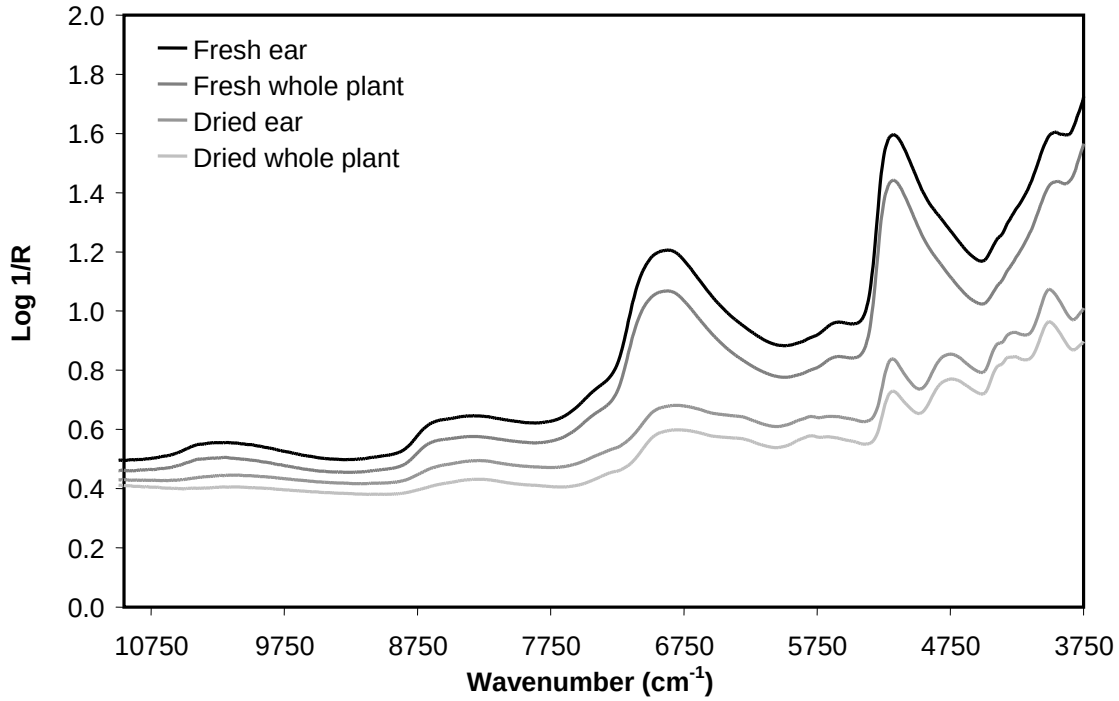
The regions which have been associated with components of cereal grains are highlighted in Figure 14 and even in the near infrared, it is clear that the presence of water results in some dominant features. In fact, it is the presence of hydroxyl groups in general which leads to the main features of the spectrum found around 7150 and 5260cm^{-1}. Both of these respond to moisture content but the absorbance region at the larger wavenumber range also contains significant levels of information about materials which contain hydroxyl groups and which are sensitive to changes in the hydrogen bonding environment of the material assessed. In addition to starch, therefore, alcohols and sugars also exhibit strong absorbances in this area of the spectrum. Starch is also responsible for the band at 4760cm^{-1} which, in wheat, dominates this part of the spectrum to such an extent that the region where protein absorbs (4590cm^{-1}) takes the form of only a slight shoulder on the starch band. This is largely due to the significantly greater concentration of starch in wheat when compared with protein and the relevant sizes of these bands changes significantly when the ratio of the components in the material assessed changes. Nevertheless this is a good illustration of the importance of mathematical pre-treatment of spectra as well as the sophisticated regression techniques which may be applied to derive NIR calibrations. Such approaches become fundamental when trying to determine protein characteristics in wheat where the absorbance band associated with protein is relatively small.

Lipids may also be determined using NIR and a band associated with this suite of components has also been marked. Again, these are relatively small in accordance with the low level of lipid generally found in wheat.

Figure 15 illustrates example spectra for each of the 4 starting materials from which NIR spectra have been acquired during the work reported. The most obvious difference between them is the dominance of water regions in the spectrum for those samples which had not experienced any drying, i.e. the fresh material. Both the immature heads and entire plant materials had moisture contents in the region of approximately 55-80% and so while the spectral information related to the other components such as protein still exist, the features associated with them are relatively smaller and incorporated in combined absorbance bands due to both water and the component of interest. This corresponds with work using NIR to determine properties of silage where the moisture content of the material measured is of the same order (Park *et al.*, 1999).

It should be noted that while the spectra for the dried materials clearly show spectral features other than simply water, the level of detail compared with that for spectra of mature wheat is still slightly reduced as the moisture content for these materials is still higher than would be typical for mature wheat grain. These example spectra serve to demonstrate, therefore, the potential difficulties which were foreseen at the outset of the work undertaken. The fresh material was felt to offer the most practical means of acquiring spectra in a relevant timescale but the spectra thus collected were dominated by the high moisture content of these materials. While the dried material had spectra which apparently gave a better starting point for calibration development, these spectra were generated using a drying process which would be expected to produce underlying changes in the molecular structure of the main components.

Figure 15. Example spectra for each of the materials for which NIR calibrations were developed



3.3.3.2 Calibration Development for Protein Content using the Entire Dataset

Calibrations were developed by using the Optimise function within Bruker's Opus software. This allows the user to set a number of parameters for calibration development such as the wavelength ranges to be assessed and the mathematical pre-treatments to be used. The software then generates calibrations for all combinations of the set parameters and ranks the results in order of their standard errors of calibration (SEC). The best of these achieved for each of the starting materials using cross-validation of the entire sample set may be found in Table 16. These results include both the performance of calibrations used for determining (predicting) the composition of the material assessed (e.g. immature ear protein from spectra taken from immature ears) as well as those used to forecast final protein content (e.g. mature grain protein content from spectra of immature ears).

It may be seen that for moisture content, the performance of those calibrations derived using undried material was better than those where drying had been used. This is because the undried material contained more moisture, providing additional information which the NIR calibration process has been able to take advantage of. In addition the material, while fresh, would be expected to be broadly equilibrated such that the surface would be expected to have a similar moisture status to that of the inside. This would be unlikely to be the case with the dried material where it would be expected that the samples would still be equilibrating when assessed by NIR. This would inevitably lead to differences between the NIR measurement and the value derived by oven drying and thus a less accurate calibration.

Table 16. Calibration performance for NIR calibrations developed using cross validation

Material	Parameter	R^2	SEC (%)	SECV (%)	RPD
Fresh ears	Moisture (%)	0.92	1.31	1.42	3.49
	Protein as is (%)	0.72	0.25	0.30	1.90
	Protein db (%)	0.78	0.52	0.61	2.11
	Harvest protein db (%)	0.51	1.18	1.42	1.43
Fresh whole plant	Moisture (%)	0.74	1.46	2.06	1.97
	Protein as is (%)	0.20	0.35	0.45	1.11
	Protein db (%)	0.72	0.97	1.07	1.90
	Harvest protein db (%)	0.41	1.24	1.54	1.30
Dried ears	Moisture (%)	0.79	1.91	2.53	2.19
	Protein as is (%)	0.83	0.32	0.42	2.48
	Protein db (%)	0.81	0.51	0.53	2.26
	Harvest protein db (%)	0.44	1.41	1.47	1.35
Dried whole plant	Moisture (%)	0.42	3.07	3.61	1.32
	Protein as is (%)	0.82	0.60	0.71	2.32
	Protein db (%)	0.82	0.75	0.83	2.39
	Harvest protein db (%)	0.60	1.13	1.27	1.57

The results for protein content expressed on an ‘as is’ basis appeared to also deteriorate when moving from fresh to dried material when assessing the values for SECV. In this case, however, the increase in error is due to the increased values for protein content in the dried material due to a proportion of the water being removed. As errors scale with the magnitude of the underlying values, this increase is predominantly related to the higher levels of protein in the dried material when expressed on an ‘as is’ basis. If the calibrations are assessed relative to the magnitude and range of the underlying values (using R^2 or RPD), then the dried samples give better performance than the fresh samples for protein content ‘as is’. The results for protein when expressed on a dry matter basis also show the same trend with the dried samples giving better overall performance. This result is due to the difference in spectral properties discussed above. As moisture is removed from the material by drying, the absorbance bands associated with components other than water become more prominent and thus more accurate calibrations may generally be derived.

For the calibrations derived for determination of the composition of the measured material, the use of whole plant material generally leads to a decrease in performance. The only exception to this is for protein content on dried material where the whole plant gives equivalent performance in terms of R^2 when expressed on both an ‘as is’ and dry matter basis. In these cases, however, the difference in range between ear and plant material has an effect as the SECV values (which are unaffected by range) are lesser for the ear material in both cases.

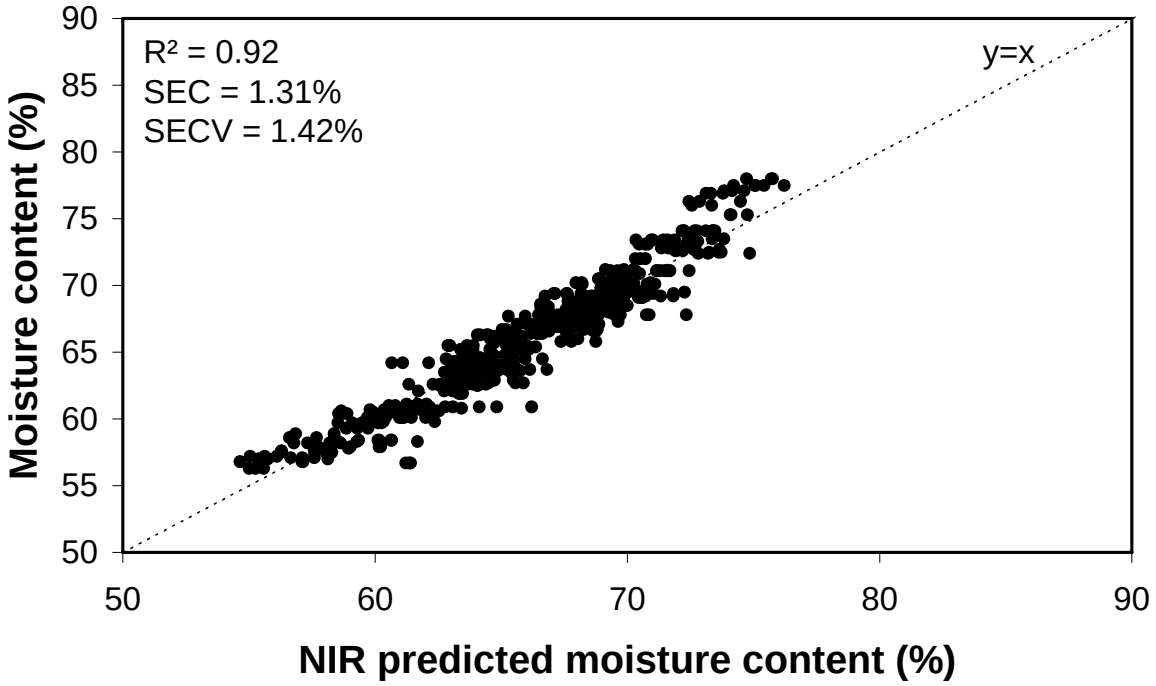
Overall, the values obtained for protein and moisture do not compare well with those that would be obtained when assessing mature grain. Targets for performance in terms of the standard error of prediction (when using a separate validation set) for ground wheat have been set at 0.3% and 0.2% for protein (as is) and moisture respectively (ICC, 1986). In practice, however, superior performance would be expected when using research grade instrumentation within a single laboratory with SECV values of 0.10% and 0.15% being reported for protein (as is) and moisture respectively in a previous study (Millar, 2003). The values obtained here for immature material are clearly poorer due to the higher moisture content and the heterogeneous nature of the starting material. Nevertheless, the results still demonstrate sufficient accuracy to allow them to be useful as a screening tool. In the absence of alternative methods to allow such measurements of protein content to be made in developing grain, therefore, the level of performance achieved would be expected to increase a user's chances of determining the crop status in a more accurate way than simple estimation.

All of the indicators of calibration performance show that forecasting of final harvest protein content when using spectra collected from immature material leads to a decrease in calibration performance when compared with determining the dry matter basis protein content directly for the material assessed. This is not unexpected given the fact that other factors would be expected to have a significant effect on the grain during the latter stages of maturation (and thus final protein content) and that these factors could not form part of the NIR calibration process. Nevertheless, the results obtained do indicate that the calibrations generated potentially have some level of predictive power. On this basis, therefore, further work assessing calibration performance using separate validation sets to reflect likely 'real-world' performance was warranted.

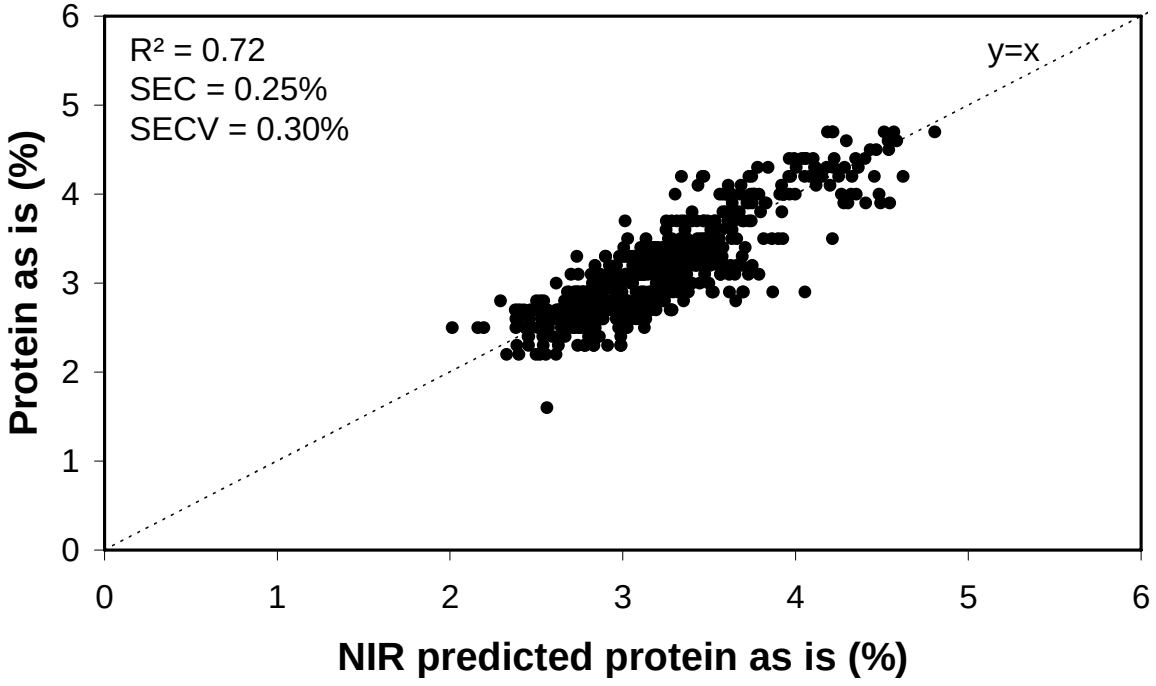
Illustrations showing examples of calibration performance may be found in Figure 16 and confirm that the calibrations fall into broadly three categories. The first represents the calibration for moisture which in the case of fresh ears gives good performance with a relatively low level of scatter around the regression line when compared with the range of values covered. Of particular interest is the fact that the distribution of points around the line of equivalence is essentially linear. This is encouraging as it would not be unexpected to find such problems occurring when very high levels of moisture were recorded (Büchmann *et al.*, 2001). The next group of calibrations contains both of those for protein content when determining the composition of the material from which NIR spectra have been collected. Finally, the forecasted results obtained when using mature grain protein values to generate a regression from immature ear spectra were the poorest with a broad distribution of points around the line of equivalence.

Figure 16. Example NIR calibrations for immature ears

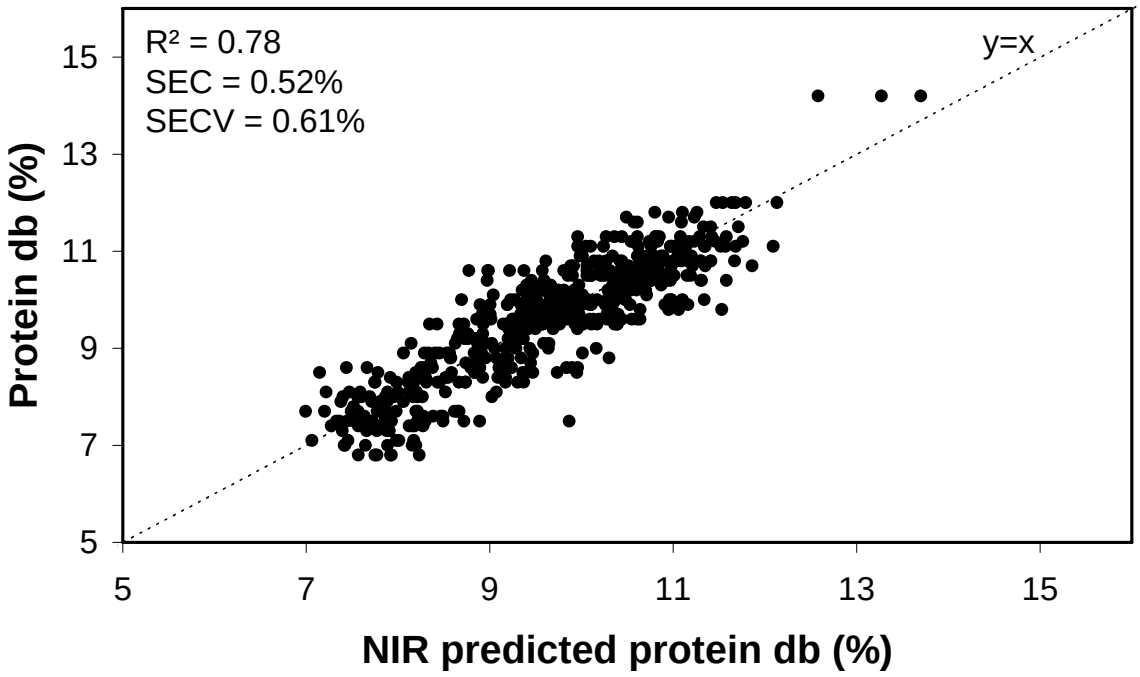
16a. Moisture content



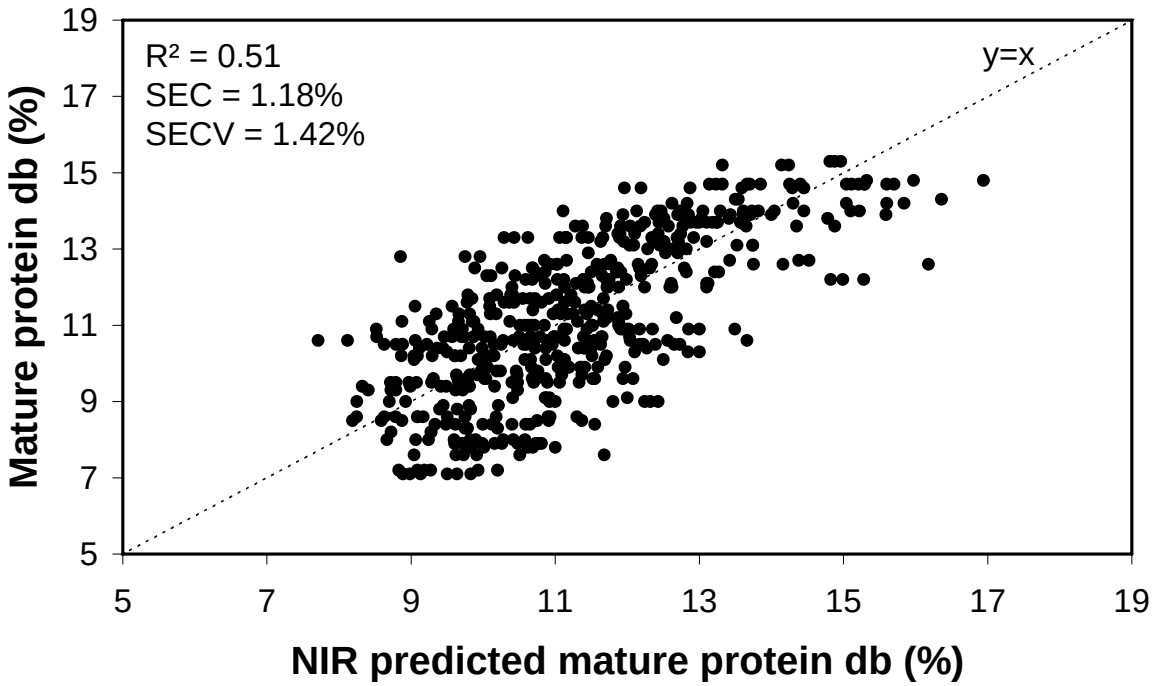
16b. Protein content as is



16c. Protein content on a dry matter basis



16d. Forecasted harvest protein



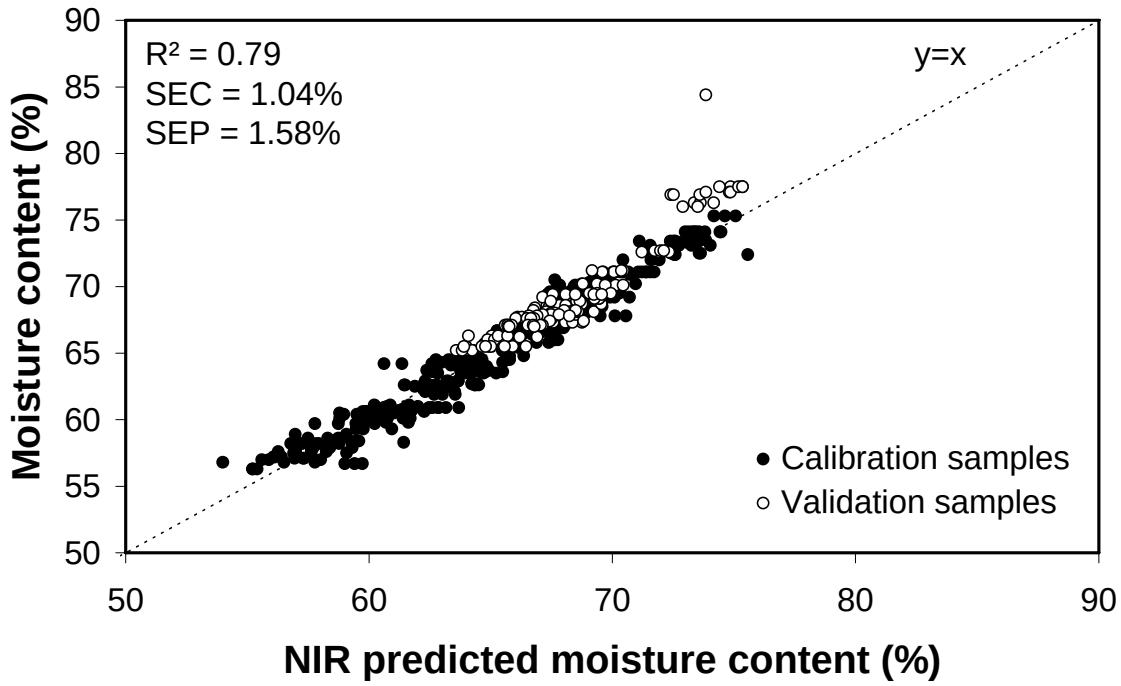
3.3.3.3 Calibration Development For Protein Content Using A Separate Validation Set

Two validation approaches were undertaken. The first of these used ear and whole plant samples grown at the Terrington site in all years as the validation set, the results of which were predicted using a calibration derived from the remaining samples for the other sites across all years. The calibration selected in each case was chosen using the Opus Optimisation software by following the same protocol as described for cross-validation. In this case, however, calibrations were optimised using the standard error of prediction rather than the standard error of cross validation. The same approach was taken for the second validation set in which all the samples (from all sites) from the 2004 harvest year were predicted using calibrations developed for all sites from the remaining harvest years (2002, 2003 and 2005).

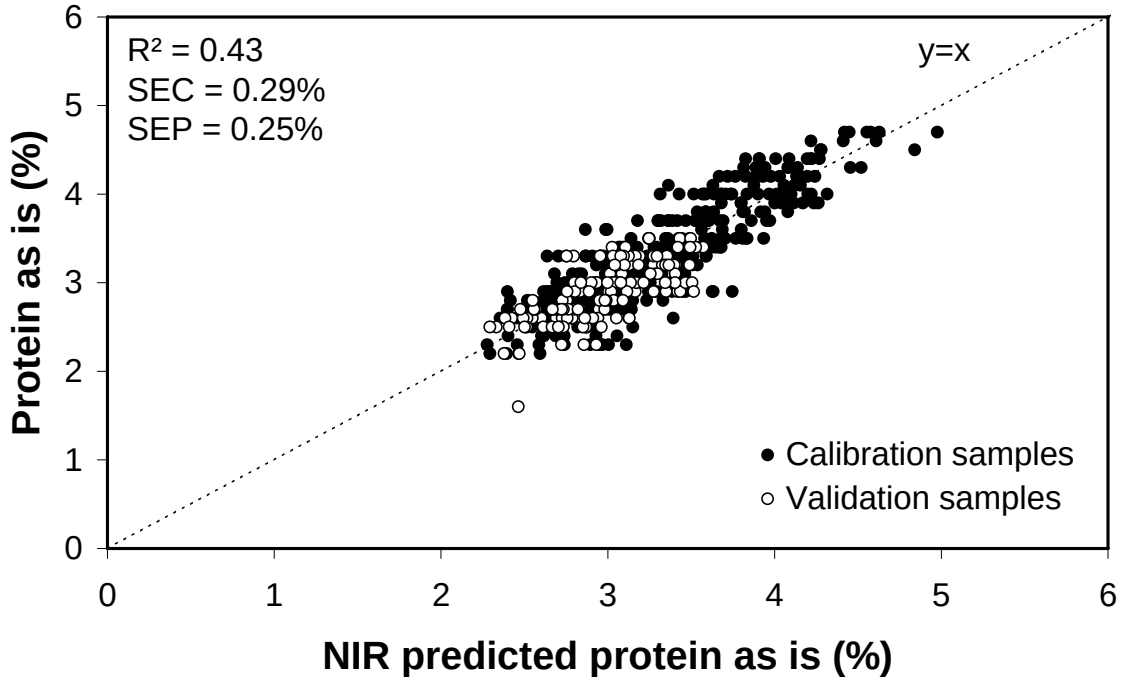
As might be expected the overall pattern of results was similar to those previously recorded when producing calibrations from the entire dataset by cross validation (Figures 17-18). Using calibrations derived from fresh ears as examples again, the three groups of calibrations (moisture, both proteins on immature material and the forecasted protein) were evident. When the results were examined in more detail (Table 17), similar trends for the different materials from which spectra were collected may be seen as those observed for the cross validation work. Calibrations for moisture content gave better performance when derived from fresh material rather than dried and, with the exception of fresh material from Terrington, ears generally gave better results than whole plants. The trends for immature plant protein 'as is' and on a dry matter basis were also consistent, with dried and head material giving lower SEPs than fresh and whole plant, respectively.

Figure 17. Example validation set performance for NIR calibrations for immature ears using data from Terrington

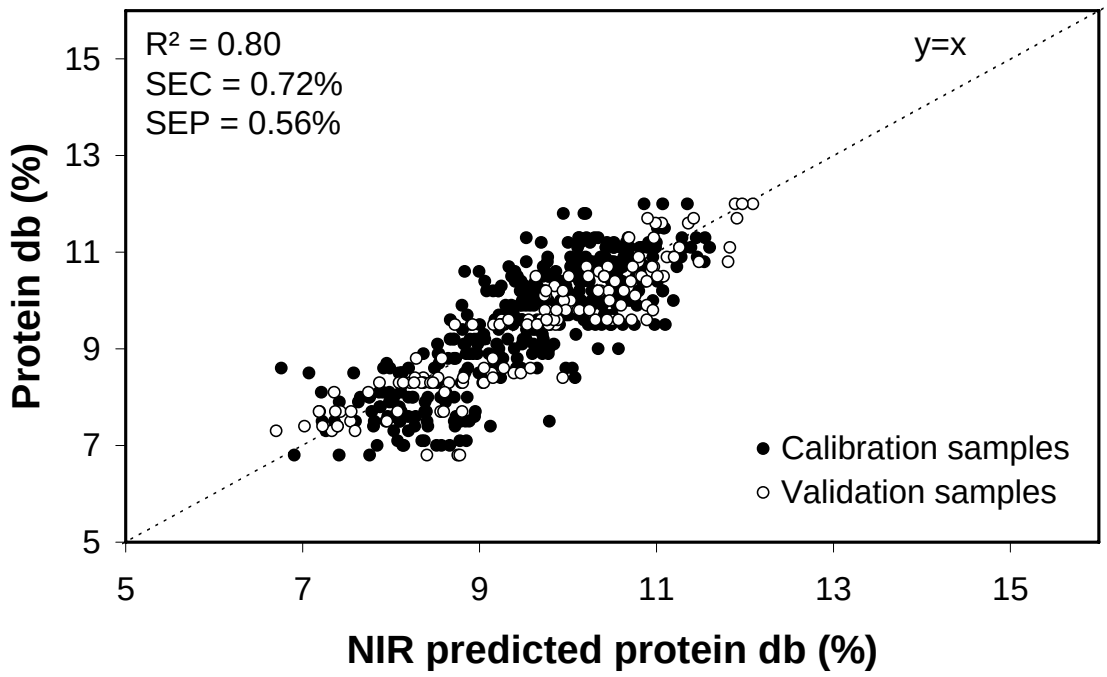
17a. Moisture content



17b. Protein content as is



17c. Protein content on a dry matter basis



17d. Forecasted harvest protein

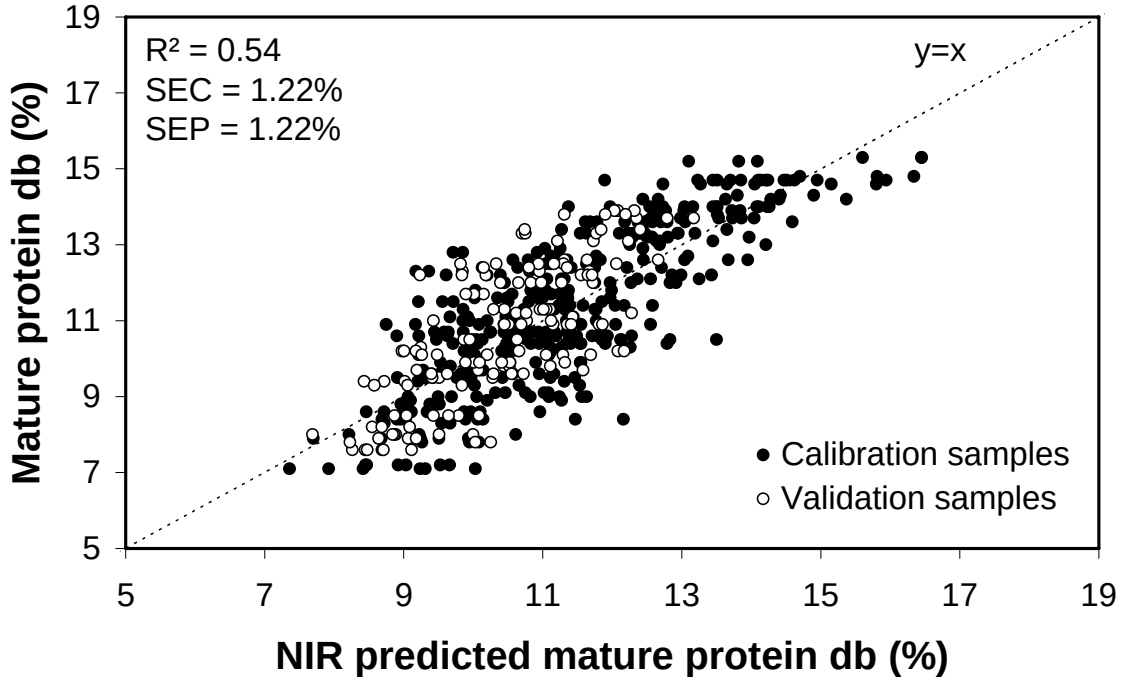
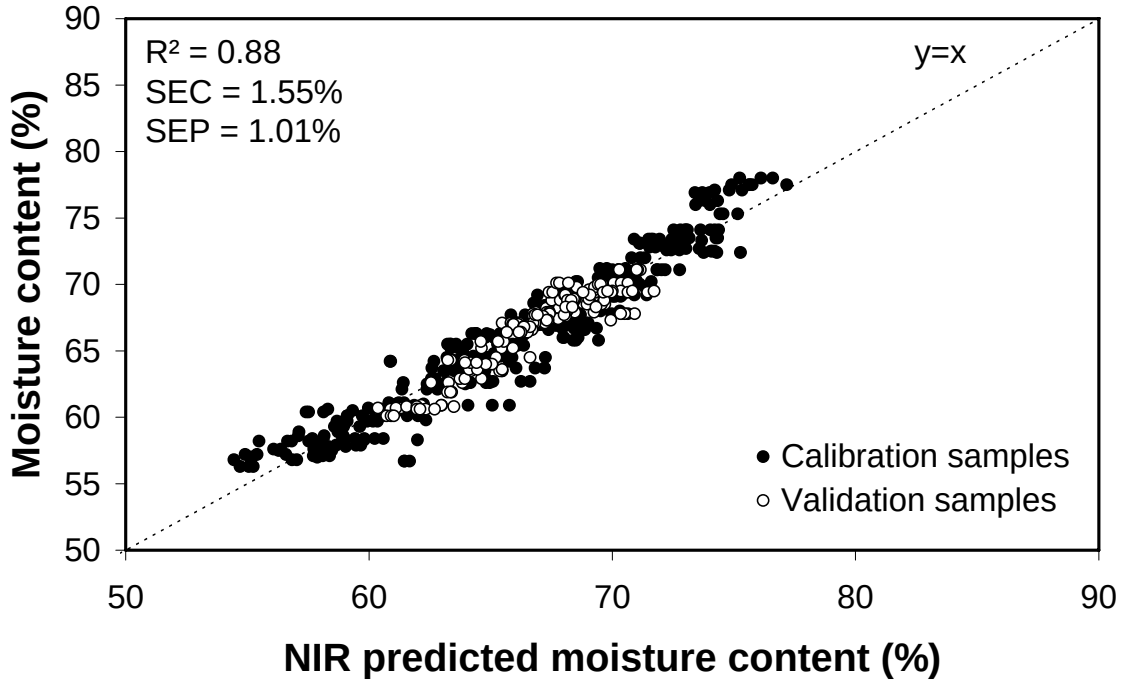
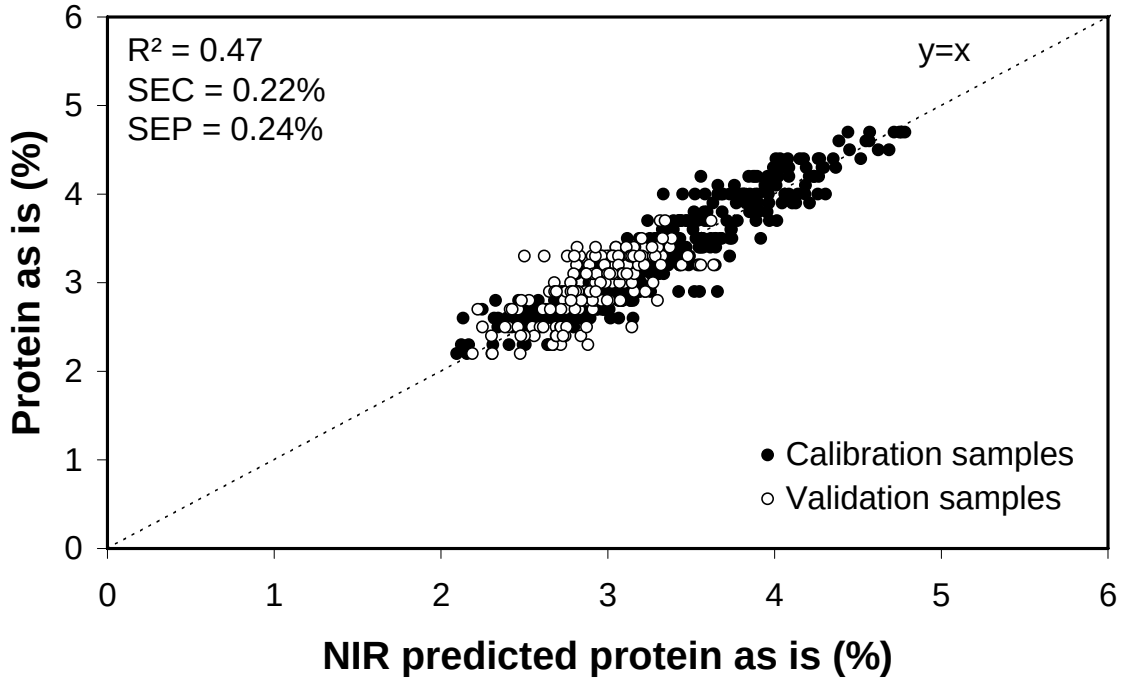


Figure 18. Example validation set performance for NIR calibrations for immature ears using data from the 2004 harvest

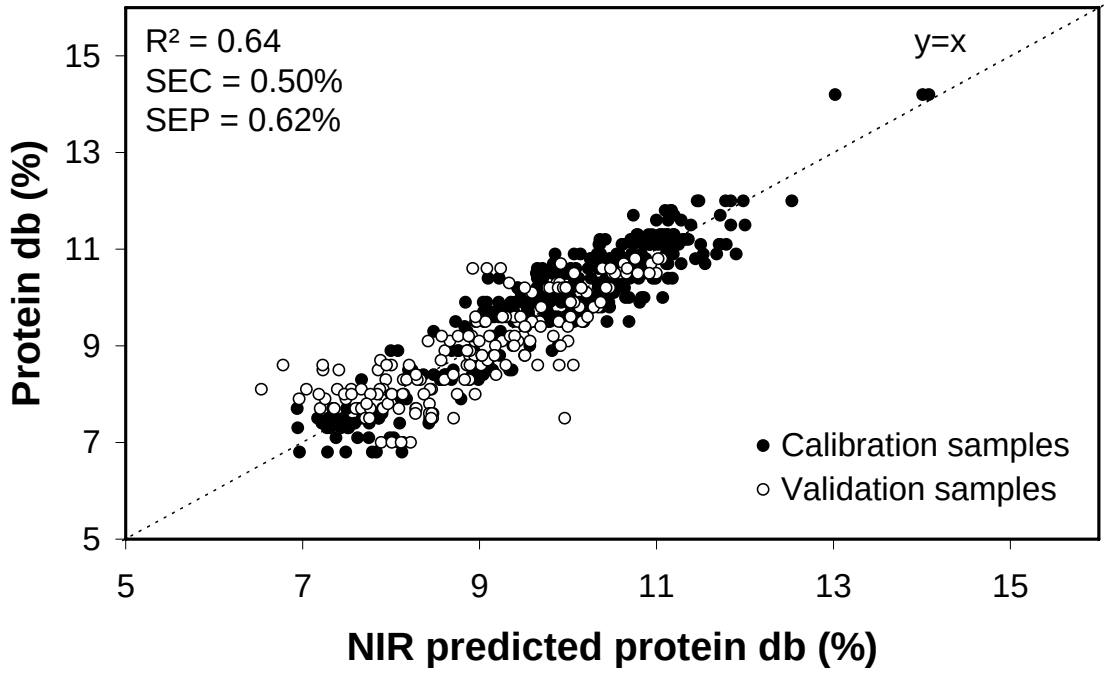
18a. Moisture content



18b. Protein content as is



18c. Protein content on a dry matter basis



18d. Forecasted harvest protein

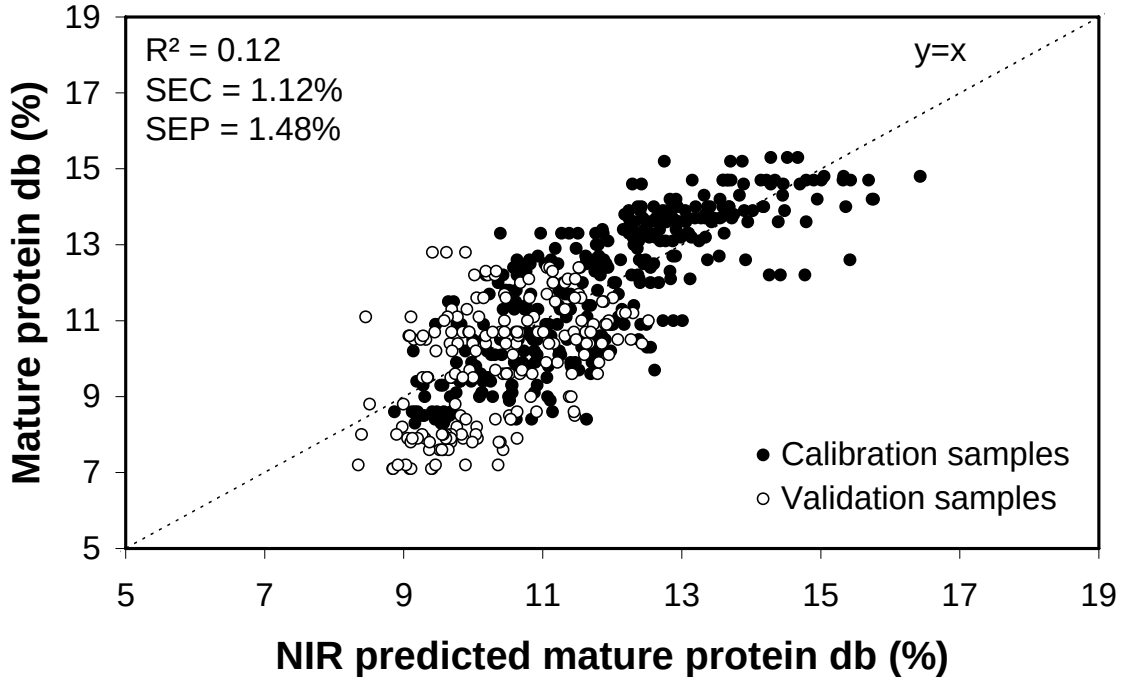


Table 17. Calibration performance assessed using a separate validation set

Material	Parameter	Terrington validation			2004 validation		
		R^2	SEC (%)	SEP (%)	R^2	SEC (%)	SEP (%)
Fresh ears	Moisture (%)	0.79	1.04	1.58	0.88	1.55	1.01
	Protein as is (%)	0.43	0.29	0.25	0.47	0.22	0.24
	Protein db (%)	0.80	0.72	0.56	0.64	0.50	0.62
	Harvest protein db (%)	0.54	1.22	1.22	0.12	1.12	1.48
Fresh whole plant	Moisture (%)	0.83	1.74	1.08	0.80	2.58	1.72
	Protein as is (%)	0.51	0.35	0.39	0.45	0.28	0.43
	Protein db (%)	0.81	1.09	0.87	0.65	1.07	1.04
	Harvest protein db (%)	0.53	1.42	1.30	0.06	1.54	1.53
Dried ears	Moisture (%)	0.73	1.85	2.36	0.82	1.95	2.36
	Protein as is (%)	0.80	0.42	0.34	0.77	0.36	0.40
	Protein db (%)	0.82	0.45	0.45	0.79	0.42	0.47
	Harvest protein db (%)	0.68	1.61	1.01	-0.11	0.89	1.67
Dried whole plant	Moisture (%)	0.42	3.47	3.43	0.23	2.74	4.52
	Protein as is (%)	0.84	0.69	0.63	0.78	0.72	0.62
	Protein db (%)	0.83	0.70	0.86	0.80	0.78	0.78
	Harvest protein db (%)	0.68	1.34	1.00	0.37	1.03	1.31

In common with the cross validation calibration work, the forecast approach to predict mature grain protein content from spectra collected from immature wheat gave poorer overall performance than the calibrations developed to determine the actual composition of immature material. However, the level of performance varied depending on the validation approach taken. In all cases (Table 17), the performance obtained when using 2004 harvest material as a validation set was poorer than that when using Terrington. This indicates that while the overall performance of the forecasting approach is generally lesser than that when dealing with composition of the actual immature material, the performance is also variable, being dependent on the validation approach undertaken. As 'real world' application of these calibrations will result in their use with a wide range of sample types, it must be concluded that the generic forecasting approach may not be recommended for such work on the basis of the results obtained here.

An additional complication which would also count against forecasting based on NIR results alone is the difficulty of taking account of variation from one harvest to another. Typically changes in NIR calibration performance with changing harvest may be accounted for by making basic mathematical changes to calibrations following an evaluation of the performance of a relatively limited set of samples immediately after harvest. These samples may either be used to adjust the calibration or be included in the overall set to improve calibration robustness. However, should such an adjustment be required for the forecasting approach, it would not be possible to apply it as the reference data (in this case the protein content of mature grain) would not be available until some months after the NIR spectra had been collected. By the time such data were available, harvest would already be over and the grower's decision making process would have been completed some time before.

For these reasons, therefore, it is concluded that the NIR approach may be used to determine the actual composition of immature grain and whole plant material. These data may then be used as an input into a modelling approach designed to account for growing conditions and likely yield of the crop in question. The lack of information of this nature when attempting to develop an 'all-in-one' NIR forecasting approach explains why the attempts at developing this sort of calibration as described above did not lead to a universally-applicable solution.

3.3.3.4 Development of Calibrations for Measures of Protein Quality

Previously NIR calibrations were developed which allowed wheat protein quality as described using the Profilblé® size exclusion high performance liquid chromatography method (SE-HPLC) to be predicted for whole grain, ground whole grain and white flour (Millar, 2003). While these calibrations were sufficiently accurate to allow them to be used to indicate protein quality (and indeed are now used for sample assessment in Recommended and National List testing), it is recognised that they do not approach the levels of performance that would be expected for determining wheat protein content by NIR. In this context, therefore, it would not be surprising to find that calibrations for protein quality derived from spectra collected from immature wheat would be poor given that the overall level of NIR calibration performance is

lesser than that expected for mature grain. Preliminary assessment of the calibrations developed for the key ratios of SE-HPLC parameters (F1, F2, F3 and F4) indicates that the R^2 values obtained were, indeed, small indicating poor correlations when using a cross validation calibration approach (Table 18). When examining the results in more detail, however, the SECVs obtained for immature wheat were slightly better than those derived for flour in the previous study (Millar, 2003). More specifically, those for immature ears appear particularly interesting. The reason for the low R^2 values is that the range of values recorded within the work reported is reduced in comparison with the previous work and, as a result, the R^2 is reduced for a given error level. The SECV results in this study are slightly better than those obtained for mature wheat and flour in the previous work and, while the calibrations derived would not be generally applicable at this stage of development, the underlying trend appears worthy of further investigation. Certainly, an underlying relationship between the ratio of gliadins to high molecular weight glutenins [(F3+F4)/F1] and protein content contributed to the calibration performance for that ratio (Table 18) but such a relationship does not exist for F1/F2, a ratio which has tended to prove more challenging for the generation of NIR calibrations. It may be that the relationship between the spectra of immature ears and their protein quality as assessed by SE-HPLC is enhanced at an early stage of development prior to grain drying, a process which would be expected to have profound implications for molecular interactions between proteins. In the absence of further work in this area, however, the underlying reasons for this interesting relationship are speculative rather than definitive.

Table 18. Calibration performance for SE-HPLC parameters assessed for NIR calibrations developed using cross validation

Material	Parameter	R^2	SEC (%)	SECV (%)	RPD	R^2 (SE-HPLC v protein db)
Fresh ears	F1/F2	0.43	0.021	0.026	1.31	0.11
Fresh ears	(F3+F4)/F1	0.45	0.18	0.24	1.38	0.30
Fresh whole plant	F1/F2	0.03	0.028	0.033	1.03	0.15
Fresh whole plant	(F3+F4)/F1	0.24	0.21	0.29	1.14	0.33
Dried ears	F1/F2	0.20	0.029	0.030	1.13	0.12
Dried ears	(F3+F4)/F1	0.28	0.25	0.28	1.18	0.31
Dried whole plant	F1/F2	0.15	0.022	0.030	1.10	0.14
Dried whole plant	(F3+F4)/F1	0.41	0.18	0.25	1.28	0.33

3.3.4 Conclusions

- Determination of the protein and moisture contents of fresh wheat ear and plant material may be performed using NIR.

- Ears generally give better performance than whole plant.

- Dried material generally gives better performance for protein but not for moisture.

- The performance of calibrations seeking to forecast grain protein at maturity was variable under validation conditions.

- Forecasting of mature grain protein content is likely to be improved by using models taking account of the state of grain development and the likely crop yield.

- Calibrations for protein quality show poorer performance than those for crude protein, on a dry matter basis, although evaluation of true performance is restricted by limited sample range.

- The error associated with NIR determination of crude protein is less than the variability of the dataset and so will give more reliable estimates than those methods currently available.

- NIR calibrations should form the basis, therefore, of an integrated crop monitoring system.

3.4 Biochemical Study of Functional Proteins in Developing Grain

Peter Shewry and Jackie Freeman, Rothamsted Research

3.4.1 Introduction

Two aspects of the protein content of cereal seeds contribute to processing quality, the total amount (protein content) and the composition and properties (protein quality). In particular, the interactions of the gluten proteins, and especially their ability to form high molecular weight glutenin polymers, are particularly important in determining grain quality.

Work at Rothamsted has therefore focused on the impact of late nitrogen application on protein quality, working in collaboration with UEA and CCFRA to define these effects and develop simplified methods to determine them prior to grain harvest. This has required the development of new methods to isolate "gluten protein" fractions from developing grain *without* altering their structures and functional properties. Effects on protein subunit composition were determined by quantitative gel scanning of total gluten protein fractions and aliquots of selected fractions were provided to UEA for analysis.

3.4.2 Materials and Methods

3.4.2.1 Preparation of Protein Fractions from Developing Grain

Protein bodies

Protein bodies were isolated from developing wheat endosperms essentially as described by Davy *et al.* (2000). Wheat kernels from 10 ears of the bread wheat variety Cadenza at about 21 days after flowering (i.e. mid-development stage) were de-husked and the pericarp was removed manually to leave only the endosperm (approximately 20g total weight). The endosperms were chopped with a razor blade in 20ml of buffer A (20mM *N*-(2-hydroxy-ethyl)piperazine-*N*'-2-ethanesulfonic acid (HEPES), pH 7.6, 100mM NaOAc, 5mM $MgCl_2$) to allow the protein bodies to float out. The sample was filtered through four layers of pre-wetted muslin, carefully layered onto 2 x 5ml 1.75M sucrose cushions (made up in buffer A) in two Corex tubes and spun in a centrifuge (500g, 2 min, 10^{o}C). The material from the top of the cushion (2mL volume) was collected and re-suspended in buffer B (20mM HEPES, pH 7.6, 100 mM NaOAc, 5mM ethylenediamine-tetraacetic acid (EDTA), 0.25M sucrose) to give a total volume of 10 ml. Two Percoll step gradients (2ml of 1.13 gcm^{-3} and then 2ml of 1.08 gcm^{-3}, in 0.25M sucrose) were prepared in 15ml Corex tubes. Half of the suspension was layered onto each of the Percoll step gradients and then centrifuged (9,000rpm, 60 min, 10^{o}C). The protein bodies were collected from the surface of the 1.13gcm^{-3} Percoll layer (2 x 500µl). The protein bodies were washed twice: 500µl of water was added, the sample was centrifuged for 2 min at 13,000 rpm and then the supernatant was removed. The pellet was stirred gently into 300µl of chloroform/methanol, 2:1, to remove membranes. After centrifugation (13,000g, 1 min) and removal of the chloroform/methanol the cohesive, elastic protein mass was stored at -20^{o}C.

Gel proteins

Endosperms (10g) were chopped in 10ml glycerol and filtered through four layers of muslin that was pre-wetted with glycerol. The filtrate was centrifuged at 41,000g for 20 min at 10°C and the pellet re-suspended in 10ml glycerol by gentle pipetting. It was then re-centrifuged and the pellet washed by gently pipetting distilled water to remove starch without providing work input into the gel protein pellet (Mutla *et al.*, 1999).

Enzymic digestion

White flour (100mg) was suspended in 1.26ml buffer (0.3M sodium phosphate: 0.5M sodium acetate, pH 6.0), or 320mg of immature seeds (pericarp removed) were chopped in 0.26ml buffer and 1ml buffer added to the suspension. 20µl β-xylanase (1851U ml^{-1}, Megazyme), 20µl endo-cellulase (713 ml^{-1}, Megazyme) and 10µl protease inhibitor cocktail (Sigma-Aldrich, UK) were added and the suspension was stirred at 60°C for 24h. 100µl α-amylase (1000U ml^{-1}, Megazyme) and 100µl amyloglucosidase (156U ml^{-1}, Sigma-Aldrich, UK) were added and the suspension stirred at 48°C for 48h. Undigested material was pelleted (13,000g for 1 min) and washed twice by re-suspension in buffer. The final pellet was stored at 4°C.

Extraction of total proteins

Homogenised endosperms (1g) or white flour (1g) was stirred with 10ml 50% (v/v) propan-1-ol, 2% (v/v) 2-mercaptoethanol, 1% (v/v) acetic acid in a falcon tube for 1h in a fume hood. The mixture was centrifuged at 10,000g for 20 minutes at 20°C. The supernatant was collected and pellet re-extracted by re-suspending in 10ml 50% (v/v) propan-1-ol, 2% (v/v) 2-mercaptoethanol, 1% (v/v) acetic acid and stirring as before, then centrifuging as before. This was repeated a third time and the three supernatants pooled. The supernatants were dialysed against 5l of 0.1M acetic acid (0.1% (v/v)) for 48h, with five changes of acetic acid solution, at 4°C. The contents of dialysis tubing were collected into a falcon tube and freeze-dried.

3.4.2.2 Determination of Carbohydrates

Anthrone determination of carbohydrate

Reagent

Concentrated sulphuric acid (74ml, specific gravity 1.84) was added slowly to 26ml of distilled water cooled with ice. The solution was cooled to about room temperature and 200mg anthrone added. The vessel was covered with foil and stirred until the solid dissolved.

Procedure

Glucose (1mM) was added in increments of 25 or 50µl to 1.5ml Microfuge tubes with screw caps, and the total volume adjusted to 200µl with water. Anthrone reagent (1ml) was added, the tubes capped, the contents mixed and transferred immediately to a heating block at 100°C. After exactly 11 min the samples were plunged into ice to cool. An aliquot of 300µl of each sample was transferred to wells of a microtitre plate and the absorbance measured at 630nm. Standards and a blank were always included with each set of

assays. A known amount (between 5 and 10mg) of the dry, enzyme digested flour samples were dispersed in 3ml of water and 200µl aliquots assayed, in the same way.

3.4.2.3 HPLC

Sugars released by acid hydrolysis of the enzyme-treated flour samples were determined by HPLC.

Hydrolysis

To 10mg of the dry, enzyme-digested flour samples was added 30µl of water, followed by 85µl of concentrated sulphuric acid (18.3M) – slowly, at 0°C, to avoid large rises in temperature. After 2h at RT, water (1.44ml) was cautiously added to adjust the concentration of sulphuric acid to 1M and the solution incubated at 100°C for 8h with occasional, vigorous, mixing. The solution was cooled, clarified by centrifugation then neutralised by prolonged agitation with a 10% excess of barium carbonate. Unreacted barium carbonate and precipitated barium sulphate was sedimented and the supernatants passed alternately through small columns (0.5ml bed volumes) of (i) a protonated, strong cation-exchanger (Dowex-50 (H^+)) to remove all traces of metal ions; and (ii) a strong anion exchanger (Dowex-1 (acetate)) to remove all traces of sulphate. The samples were then dried by evaporation *in vacuo,* over anhydrous calcium chloride and sodium hydroxide pellets. The residues were rehydrated in 250µl water and frozen. Aliquots of 2.5-10µl were analysed for total carbohydrate, using the anthrone method (above) and the remainder kept for HPLC analysis.

HPLC

Sugar analysis was achieved by high-performance anion-exchange chromatography with pulsed amperometric detection, using a Dionex DX-500 chromatography system, consisting of a GP40 gradient pump, an AS3500 autosampler and an ED40 electrochemical detector connected to an amperometry cell with a gold working electrode. Separation was achieved using a CarboPac PA 10 analytical (4 x 250mm) column with 18mM NaOH as eluent at a flow rate of 1ml min^{-1} throughout. Twenty minutes after sample injection, the column was regenerated using 200mM NaOH for 10 min and 18mM NaOH for 20 min, after which the next sample was loaded. Standard sugars included arabinose, glucosamine, galactose, glucose, mannose, xylose and ribose. All gave near linear responses for loadings of 0-1nmol. Processed hydrolysates were diluted 100-fold in water, filtered, and 5–20µl injected.

3.4.2.4 SDS-PAGE and Gel Scanning

Total protein fractions were extracted from single seeds and separated by SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) as described by Shewry *et al.* (1995), using a statistically randomised design for loading. The extraction buffer (0.0625M tris-HCl, pH 6.8, 2% (w/v) SDS, 5% (v/v) 2-mercaptoethanol, 10% (v/v) glycerol, 0.002% (w/v) bromophenol blue) was also used to dissolve protein preparations for SDS-PAGE. Gels were stained in 0.1% (w/v) Coomassie Brilliant Blue R250 in 40% (v/v) methanol, 10% (w/v) trichloroacetic acid and destained in 10% (w/v) trichloroacetic acid.

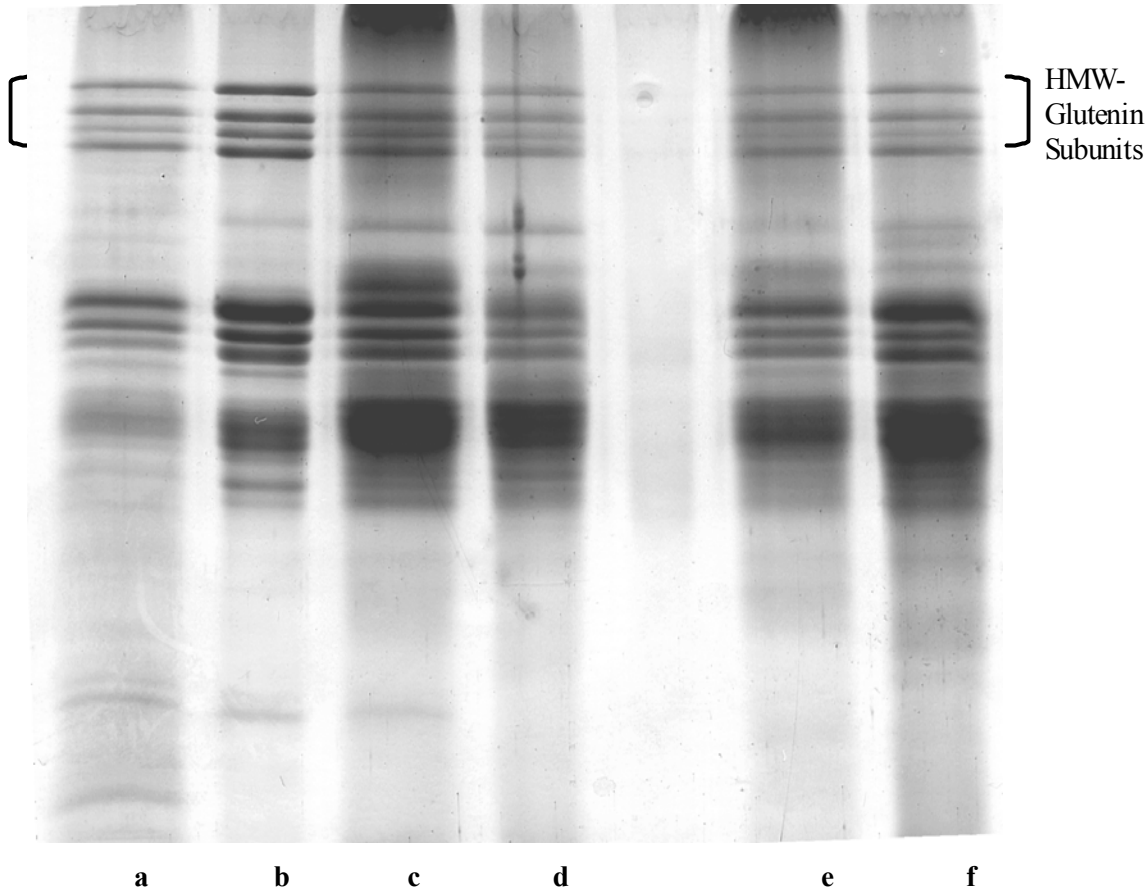
Scans of gels were analysed using *Phoretix*[tm] 1D Advanced software (Nonlinear Dynamics, Newcastle, UK), with three replicate samples separated on two gels run on different days. Means and SEMs of the six scans are presented.

3.4.3 Results and Discussion

Evaluation of methods for preparation of gluten protein fractions from developing and mature grain

Three methods were evaluated for the preparation of "native" gluten protein fractions from developing grain (at 21 days after anthesis, daa) and flour from mature grain of Cadenza, comparing the compositions of the fractions with "total" gluten protein extracted from the whole tissue and flour. The methods were initially tested on 21 day old developing grain using SDS-PAGE to monitor the composition and in particular the proportion of HMW-G subunits (Figure 19, Table 19).

Figure 19. SDS-PAGE of gluten protein fractions prepared from Cadenza using a range of methods



a. gel protein from 21 daa endosperms
b. protein bodies from 21 daa endosperms
c. enzyme digest of 21 daa endosperms
d. total proteins from 21 daa endosperms
e. enzyme digestion of flour
f. total proteins from flour

Table 19. Determination of the % HMW-G subunits in gluten protein fractions prepared from developing and mature grain of Cadenza

Sample	Method	% HMW–G subunits	
		Mean	SEM (±)
21 daa	gel protein	23.57	0.657
	protein bodies	24.54	1.208
	enzymic digestion	7.10	0.411
	total	10.97	0.302
flour	enzymic digestion	8.33	0.541
	total	11.89	0.212

Based on the means of two replicate scans each of three replicate preparations

The gel protein fraction was prepared by chopping developing endosperms in glycerol. This fraction forms a cohesive mass with properties broadly similar to those of gluten, but contains substantial amounts of starch and therefore may not be suitable for mechanical or spectroscopic studies (Section 3.5). SDS-PAGE and gel scanning showed that it was substantially enriched in HMW-G subunits (approximately 24% of the total protein) compared with the total fraction (approximately 11%), indicating that it is also enriched in glutenin polymers. It is probable that this enrichment results from loss of gliadin monomers during isolation, and thus is not a good representative of gluten.

The protein body fraction prepared from the same tissue was substantially free of starch but showed a similar level of enrichment in HMW-G subunits (to approximately 25% of the total fraction). In this case, the enrichment may have resulted from the preferential isolation of glutenin-enriched protein bodies derived from the endoplasmic reticulum as opposed to gliadin-enriched protein bodies derived from the vacuole. The latter are likely to be more fragile and disrupted during preparation. Although protein bodies are suitable for spectroscopy and mechanical measurements (Wellner *et al.*, 2005) their preparation is time-consuming with high variation in yield between replicate preparations (results not shown).

Because the gel protein and protein body preparations are not suitable for routine analysis, a third procedure was developed, in which the gluten protein was released by digesting slices of 21 day old endosperms with hydrolases to remove starch and cell wall polysaccharides. Analysis of triplicate fractions using the anthrone method gave 4.9% ± 0.7% sugars. Further analysis of the same fractions by HPLC showed 6.05% ±0.94% total sugars after hydrolysis, comprising about 87% glucose (presumably derived predominately from starch) with the remainder being arabinose, galactose and mannose or xylose (presumably derived from arabinoxylan, arabinogalactan and glucomannam polymers in the cell walls). SDS-PAGE showed that the fraction contained about 7% HMW-G subunits compared with about 11% in the total fraction. Similar values were obtained when the procedure was repeated on flour from mature grain, with 8% HMW-G subunits compared with 12% in the total fraction. Finally, the enzymic digestion method was also used to

compare fractions from developing grain of Cadenza. This showed similar patterns about 7-8% HMW-G subunits at all stages of development, from 14 to 42 days after flowering (Table 20).

Table 20. The proportions of hmw-g subunits in gluten protein fractions prepared from developing grain of cadenza by enzymic digestion

Days after anthesis	**% HMW subunits**	
	Mean	SEM (±)
14	7.29	0.529
18	7.23	0.350
22	7.50	0.406
26	8.10	0.441
30	7.04	0.378
42	8.23	0.527

Results are the mean of two replicate gel scans each of three replicate preparations

The values for % HMW-G subunits in fractions prepared using enzymic digestion were in all cases slightly lower than those in the total fractions and also lower than those reported previously (generally in the range 10-12%, see Halford *et al*., 1992; Seilmeier *et al*., 1991). It is considered unlikely that substantial amounts of glutenin-enriched polymers would have been lost during preparation and there is no simple explanation for this difference. However, it is clear that the enzymic digestion allows the isolation of substantially pure gluten protein fractions from developing and mature grain without mixing or shearing.

Analysis of Xi19 and Riband

Samples of developing (around 21 days after anthesis) and mature grain of Riband and Xi19 were obtained from the field trials carried out at ADAS Boxworth in 2004/5, from plots grown with low, medium and high levels of ammonium nitrate fertiliser input.

Total proteins were extracted from the frozen developing grain and milled mature grain and SDS-PAGE used to determine changes in the proportions of HMW-G subunits (and hence presumably glutenin polymers). The data are reported in Table 21.

Table 21. The proportions of HMW-G subunits in total protein fractions from developing (approximately 21 days after anthesis) and mature grain of Xi19 and Riband grown at ADAS, Boxworth in 2004/05

Cultivar	Age	N fertiliser	% HMW-G subunits	
			Mean	SEM (±)
Xi19	21 daa	low	9.89	0.418
	21 daa	medium	8.74	0.689
	21 daa	high	8.77	0.776
	mature	low	12.51	0.562
	mature	medium	13.49	0.754
	mature	high	13.30	0.579
Riband	21 daa	low	5.79	0.250
	21 daa	medium	5.75	0.370
	21 daa	high	5.90	0.476
	mature	low	11.81	0.515
	mature	medium	11.82	0.147
	mature	high	11.91*	0.810

Results are based on the analysis of two replicate gel scans each of three replicate preparations except * where only one replicate preparation was analysed

The results show that the proportions of HMW-G subunits increased in both varieties between approximately 21 days after anthesis and maturity. However, proportions were greater at both stages in Xi19 while Riband showed greater increases between the two stages.

No relationship between N application and % HMW-G subunits was observed, either at maturity or at 21 days after anthesis.

3.4.4 Conclusions

- A novel method has been developed to prepare "native" gluten protein fractions from developing and mature wheat by digestion with carbohydrases to remove starch and cell wall polysaccharides.

- Quantitative SDS-PAGE of total protein fractions extracted from developing and mature grain from the field trials at ADAS, Boxworth showed that the proportions of HMW-G subunits varied depending on stage of development and variety but were not affected by the N fertiliser level.

3.5 Spectroscopic Study of Gluten Proteins

DMR Georget and PS Belton, University of East Anglia

3.5.1 Introduction

Amongst the many wheat grain constituents, gluten proteins are regarded to be the predominant components governing the end quality in bread manufacture. Genetic factors and environmental growth conditions influence the nature and the properties of gluten proteins as quality determinants. The effects of N fertiliser application, in particular, have been studied previously (Luo *et al.*, 2000; Ruske *et al.*, 2004). Several workers have shown that gliadin to glutenin ratios change as a consequence of applying N as a fertiliser. Wieser and Seilmeier (1998) showed that this ratio increases when the N fertiliser input increases. More recently, Johansson *et al.* (2004) have reported that gluten 'strength' (rheological parameter) can be lowered through N fertiliser application. In the present study, we sought to investigate how the application of varying levels of N fertiliser influences gluten quality, using high resolution spectroscopic techniques which are complementary to some of the conventional biochemical methods used for characterising proteins.

During the last 15 years, infrared (IR) spectroscopy has proved to be a powerful tool for providing molecular details of structural changes within proteins that influence their functionality. Indeed, studies from Wellner *et al.* (2005), Georget and Belton (2006) and van Velzen *et al.* (2003) showed that IR spectroscopy can reveal changes in the secondary structure in gluten, e.g. from β-turns into β-sheets, following extension of the protein polymer network. It is believed that specific conformational changes such as those described by "the loop and train" model for glutenin structure can partly explain some of the processes that occur during dough mixing and development stages of the breadmaking process (Belton, 2005). In addition to FT-IR (Fourier transform infrared), this study has utilised FT-Raman (another non-invasive method) to investigate the properties of native, "unworked" gluten proteins that have been prepared in a number of different ways, including those described in Section 3.4.

3.5.2 Materials and Methods

Four forms of samples were investigated:

a) **Thin sections of immature grain** samples of Riband and Xi19 crop treated with early N fertiliser at rates of 0, 110, 190 and 270kg N/ha. These were taken from ear samples that had been microwave-dried by CCFRA and stored under vacuum at –20° C and were used for FT-IR analyses. Effects of soaking cut sections in de-ionised water for 1h were studied.

b) **Gel protein** preparations made by Rothamsted Research using developing grains of Riband and Xi19, grown at Boxworth with early N fertiliser treatments of 40, 200 and 300kg N/ha. See Section 3.4 for details of the preparation method. FT-IR analyses were performed on these samples.

c) **Protein body** preparations made by Rothamsted Research using developing grains of Riband and Xi19, grown at Boxworth with early N fertiliser treatments of 0, 150 and 270kg N/ha. See Section 3.4 for details of the preparation method. FT-IR analyses were performed on these samples.

d) **Flour** samples from Hereward, Einstein, Option and Xi19 harvest wheat crops, grown at Boxworth with early N fertiliser treatments of 40, 80, 100 and 200kg N/ha. FT-IR and FT-Raman analyses were performed on these samples.

3.5.2.1 Techniques

FT-IR spectra were acquired using a Bio-Rad FTS165 spectrometer equipped with an ATR (Attenuated Total Reflectance) Golden Gate accessory. The spectral resolution was 2cm^{-1} and the number of scans was 20. The spectra were corrected for water and water vapour and analysed using Omnic v6.1 software (Thermo Nicolet Corporation, Madison, USA). Fourier self deconvolution (FSD) was carried out in some of the experiments, using an enhancement factor of 1.3 and bandwidth of 30. The precise positions of the absorbance peaks located within the amide I region were determined using the second derivative. Intensities at different wavenumbers were normalized to the glutamine absorbance peak taken at 1600 cm^{-1} or to the random coils and □helices absorbance peaks taken at 1644 cm^{-1}.

Raman spectra were recorded using a Bruker IFS 33 FT-IR spectrometer equipped with a FRA 106 Raman module and a liquid nitrogen-cooled germanium detector. A diode laser operating at 1064nm was used as the excitation source, with ~100mW of laser power at the sample. The spectral resolution was 4 cm^{-1} and each spectrum was an average of 500 or 1000 scans. The samples were held at ambient temperature and the spectra were analysed using Omnic v6.1A software. Signal normalisation was carried out with the ferulic acid absorbance peak at 1600 cm^{-1}.

3.5.3 Results and Discussion

3.5.3.1 FT-IR Spectroscopy

Inspection of the series of spectra taken from the thin sections of immature grain samples revealed a high level of in-sample variation. Despite the care taken in preparing grain slices, the resultant scans were not reproducible. Therefore, no conclusions could be drawn about the effect of either variety or of N treatment from these data. Figure 20 shows a typical FT-IR spectrum of a thin section of Xi19 grain, containing the assigned amide I and amide II regions.

Figure 20. FT-IR spectrum (amide I and amide II regions) of a thin section of a dried immature grain sample of Xi19

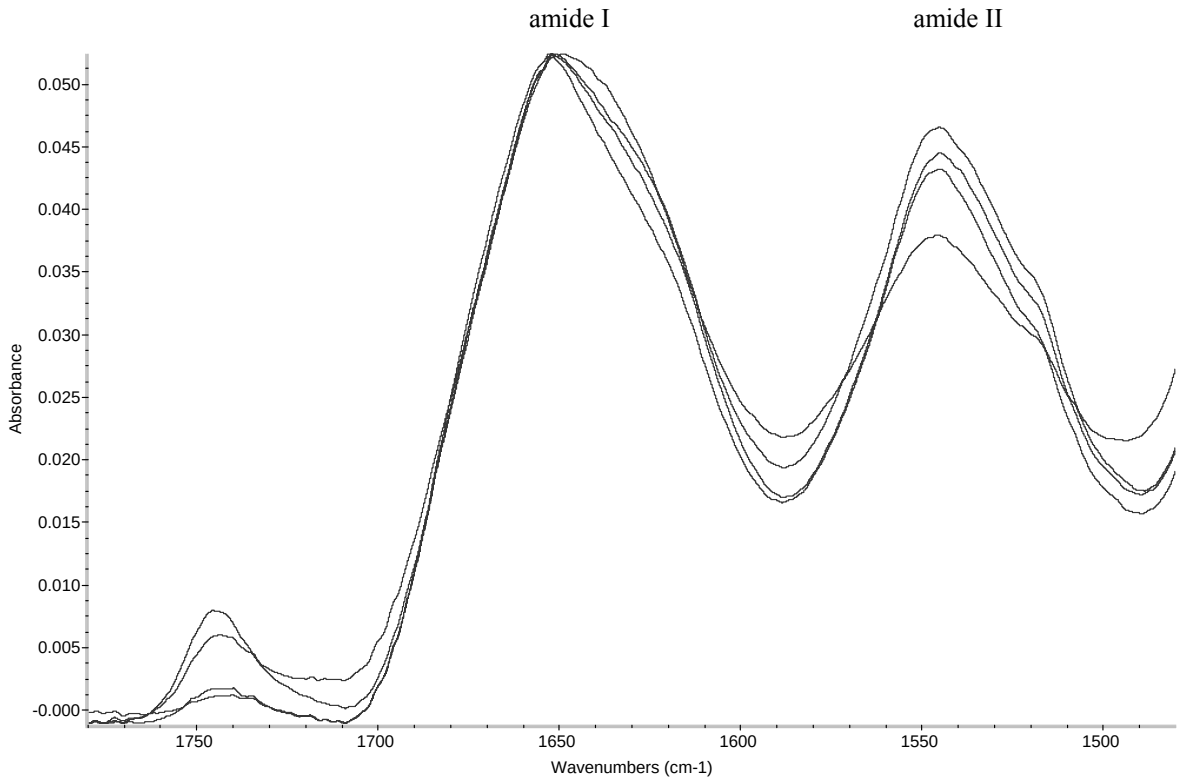
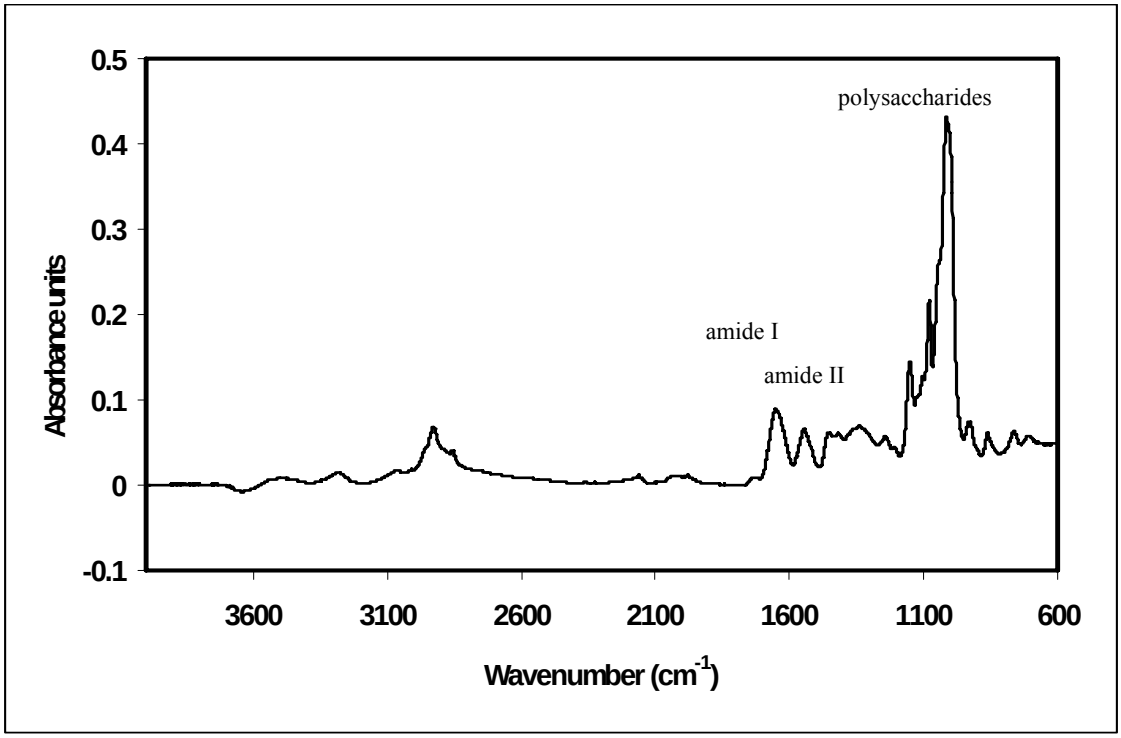
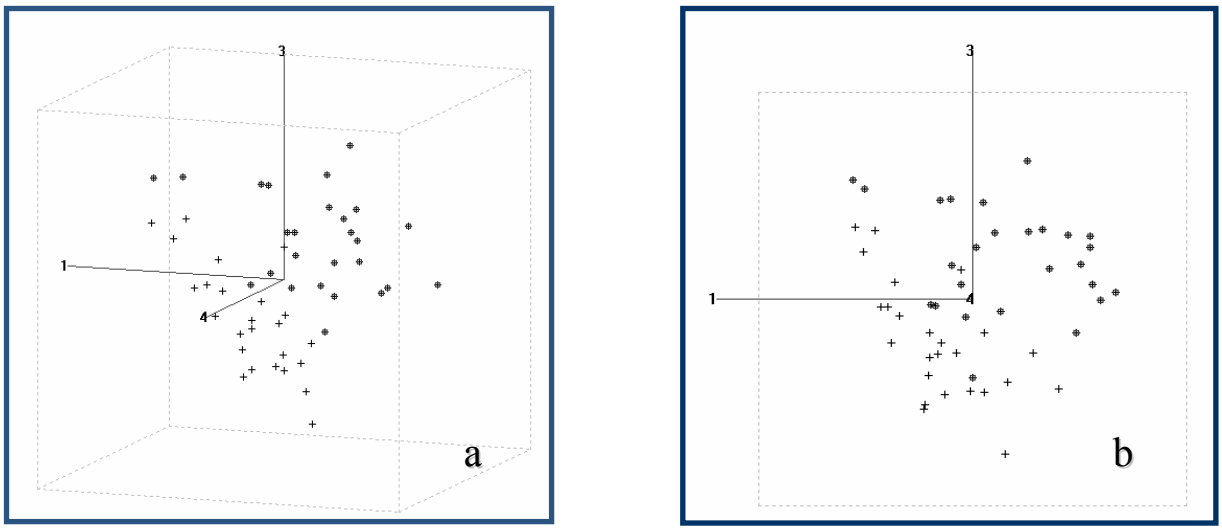


Figure 21. A typical FT-IR spectrum of a gel protein sample showing the amide I, amide II and polysaccharide regions

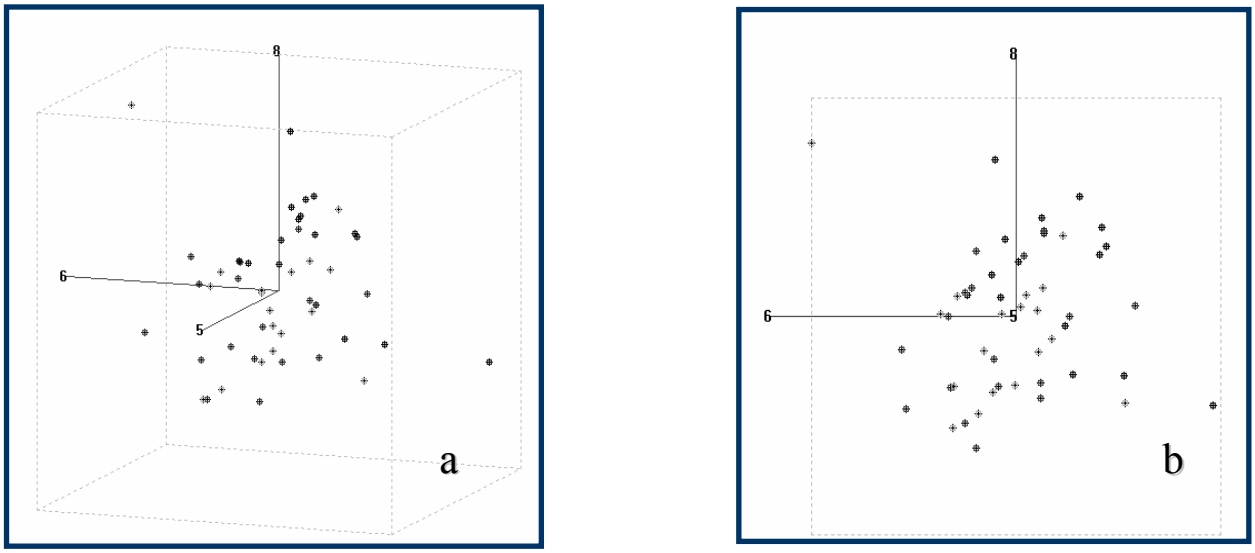


The spectra taken of the gel protein preparations contained features that were characteristic of random structures, β-turns, β-sheets and α-helices. Figure 21 illustrates a typical FT-IR scan of a gel protein sample in which the predominant peaks are: polysaccharides at around 1000 cm^{-1}; the amide II band at 1550 cm^{-1}; amide I region at 1650 cm^{-1}. Other features that were also identified included CH, CH_2 and CH_3 vibrations at around 2900 cm^{-1}; OH and NH vibrations at 3200 cm^{-1}. Comparison of the replicate scans (triplicates) showed poor reproducibility. It is most likely that this due to the highly heterogeneous nature of these samples.

Figure 22. PCA of FT-IR spectra (amide I) of gel protein preparations from immature grains. Effect of variety; dots: Xi19, crosses: Riband



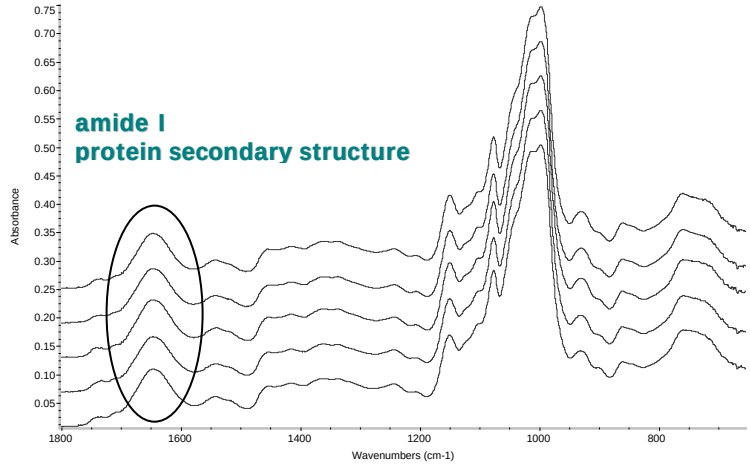
Figures 23: PCA of FT-IR spectra (amide II) of gel protein preparations from immature grains. Effect of N treatment; black- 40kg N/ha, medium grey- 200kg N/ha, light grey- 300kg N/ha



Principle Components Analysis (PCA) was performed on these spectra (by Sam Millar, CCFRA). Consequently, it was possible to separate the two varieties based on their protein absorbance values within the amide I region of the FT-IR spectra (Figure 22). It was not possible to achieve separation based on different N fertiliser treatments (Figure 23). PCA has been previously reported for NIR spectra from wheat proteins (Bhandari, 2000) and from barley seed (Jacobsen *et al.*, 2005). Blanch *et al.* (2003) have employed PCA to distinguish α-gliadins from ω-gliadins based on vibrational Raman optical activity spectra. These authors observed that the β-gliadins had more disordered structures and fewer α-helices than the α-gliadin fractions. In their study, the authors used purified and isolated materials, whereas in the present study, the material under investigation is significantly more complex.

Inspection of the replicate FT-IR spectra taken from the various protein body preparations from immature grain material showed that they lacked reproducibility (results not shown) and it was not possible to draw any conclusions. The flour samples obtained from Einstein, Hereward, Option and Xi19 harvest wheat yielded highly reproducible spectra, being more homogenous than either the gel protein or the protein body preparations. Figure 24 shows five replicate FT-IR spectra for flour from Xi19, grown with 80kg N/ha. The amide I and amide II are detected at 1650 and 1550 cm^{-1}, respectively. There is a group of multiple absorbance peaks located at around 1000 cm^{-1}. This feature is characteristic of polysaccharides. Starch components (amylose and amylopectin) and cell wall arabinoxylans contribute to these absorbances.

Figure 24. Replicate FT-IR spectra of Xi19 flour samples from wheat grown with 80kg N/ha of fertiliser.



The amide I region is commonly used to determine the changes in the secondary structure of proteins. The major absorbance peaks are shown in Table 22. The intensity values were calculated relative to the 1650 cm^{-1} absorbance peak (Figure 25) and the 1598 cm^{-1} glutamine peak (not shown) for the two sets of wheat flour samples. The error bars shown in Figure 25 are the values of standard error of the mean. It was found that this set of data were more reproducible.

Table 22. Assignment of the different amide I absorbance peaks in FT-IR spectra

Wavenumber cm^{-1}	Secondary structure feature
1699	β-turns or β-hairpins
1684	β-sheets
1670	β-turns
1650	Random coils and α-helices
1629	Anti-parallel β-sheets; more weakly hydrogen-bonded β-sheets
1614	Strongly hydrogen-bonded β-sheets; β-edges; possible contribution from glutamine side chains; intra-molecular β-sheets
1598	Glutamine side chain ($-NH_2$)

The Student's t-Test on the data obtained from the peaks in the β-sheets region revealed a significant difference between the relative peak ratios at approximately 1614 cm^{-1}, corresponding to low and high level N fertiliser treatments for the Einstein and Xi19 samples, but not for the Hereward and Option samples. These results, when normalised either to the glutamine absorbance peak at 1600 cm^{-1} or to the random coil and α-helices absorbance peak at 1644 cm^{-1}, are in good agreement with the report published by Wetzel *et al.* (2003).

3.5.3.2 FT-Raman Spectroscopy

The FT-Raman spectral analysis was performed on flour samples obtained from Einstein, Hereward, Option and Xi19 harvest wheat grown at Boxworth. Figure 26 shows typical FT-Raman spectra (5 replicates) for Option flour from crop that was treated with a fertiliser rate of 100kg N/ha. The scan profiles were similar for the five replicates. However, the signal to noise ratio was lower than expected. Figure 27 and 28 depict the relative peak intensities (normalised to the second ferulic acid absorbance peak at 1600 cm^{-1} on the basis that ferulic acid levels are believed to be unaffected by N fertiliser treatment). In the case of Option (Figure 27) there were decreases in signals assigned to phospholipids, arabinoxylans (1126, 1107 and 1081 cm^{-1}) starch backbone (940 and 478 cm^{-1}), tyrosine (865 cm^{-1}) and S-S (523 cm^{-1}) with increasing N input. This result is consistent with findings reported by Williams *et al.* (1994). Their study showed that when wheat plants were treated with high concentrations of N fertiliser fewer lipids were accumulated in the grains compared to the control samples. In this present study, the lowered lipid, arabinoxylan, starch, tyrosine and S-S levels were seen only in Option and not in the other 3 varieties. However, it should be noted these observations were based on the assumption that the ferulic acid content was the same in all 4 varieties.

Figure 25. FT-IR intensity ratios (relative to the absorbance at 1644 cm^{-1}, taken from 5 replicates) of flour samples of Xi19, Hereward, Option and Einstein

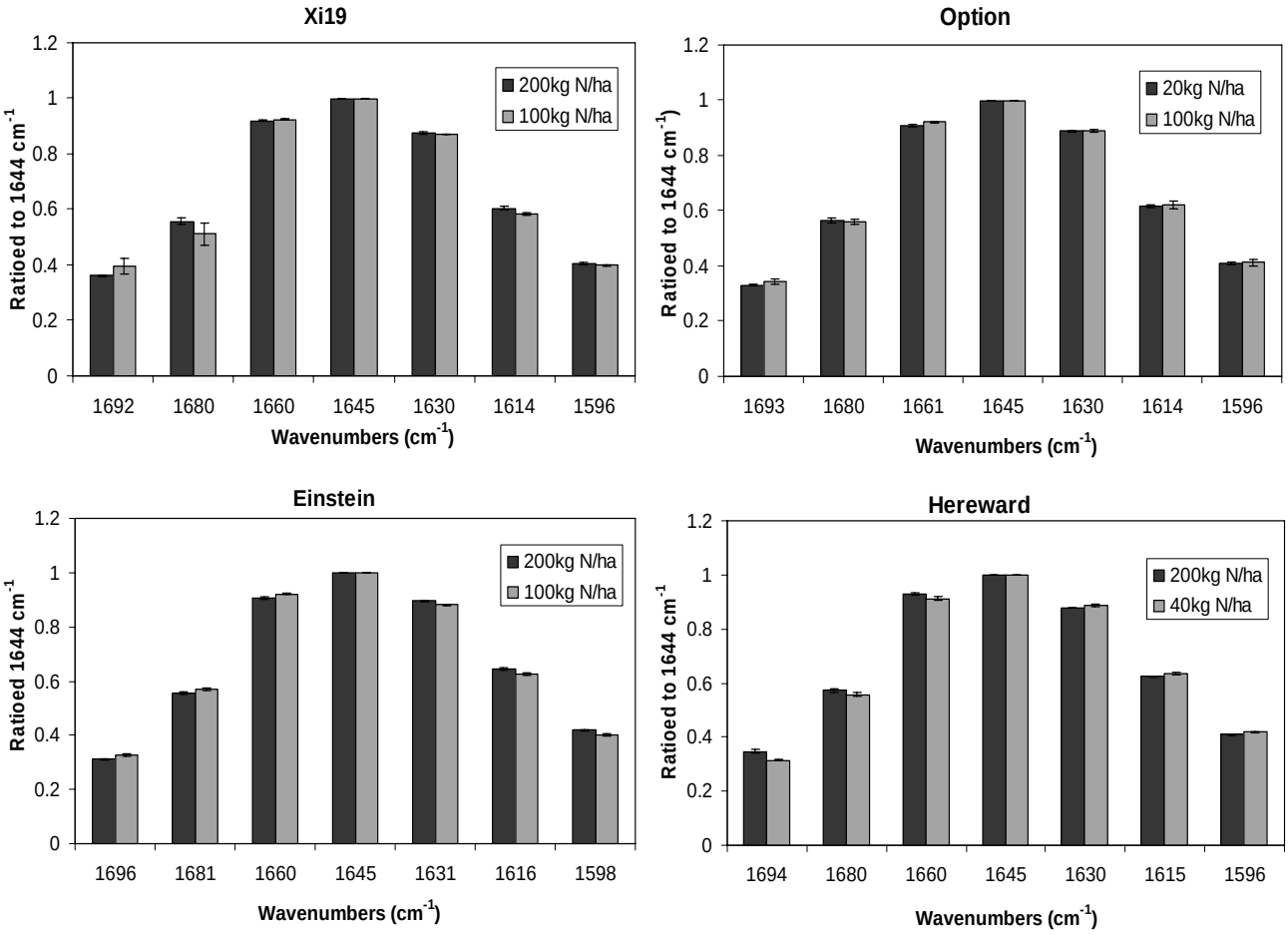


Figure 26. FT-Raman spectrum of an Option flour from wheat grown with fertiliser rate of 100kg N/ha. Spectral assignments were according to Piot *et al.* (2000 & 2001); Ram *et al.* (2003), Thygesen *et al.* (2003); Ma and Phillips (2002)

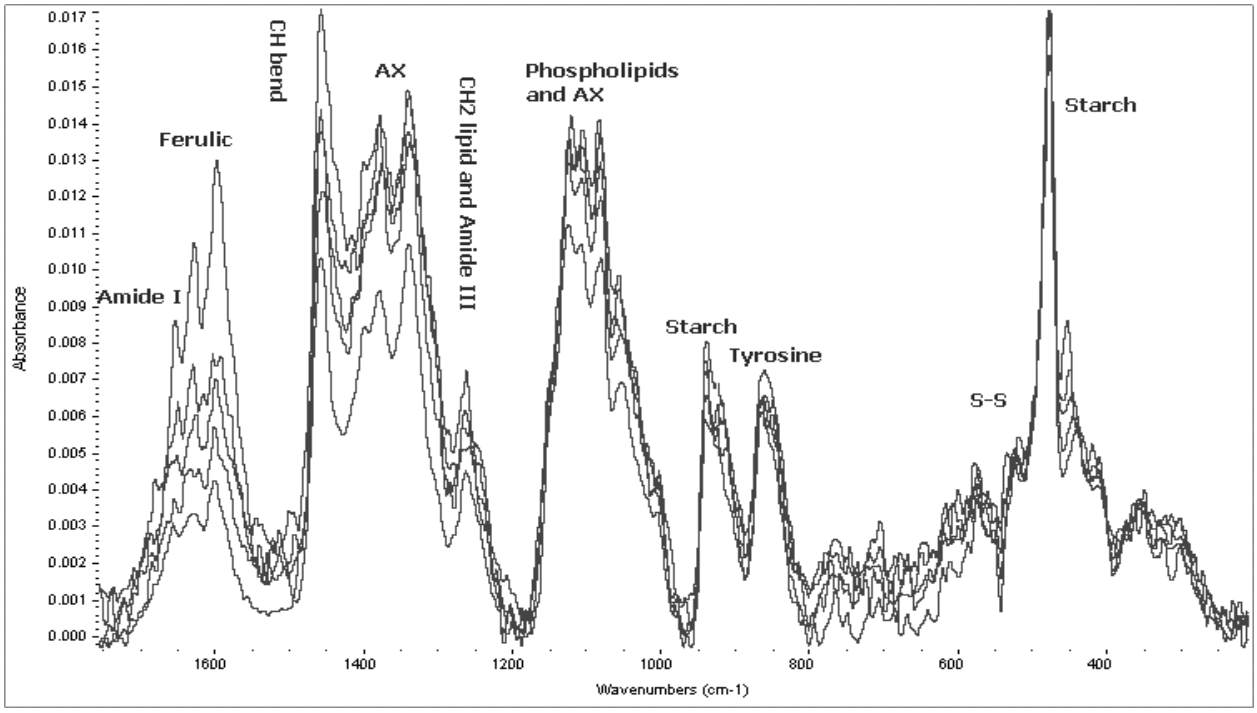


Figure 27. FT-Raman of Option flour from wheat grown with fertiliser rates of 100kg N/ha and 200kg N/ha

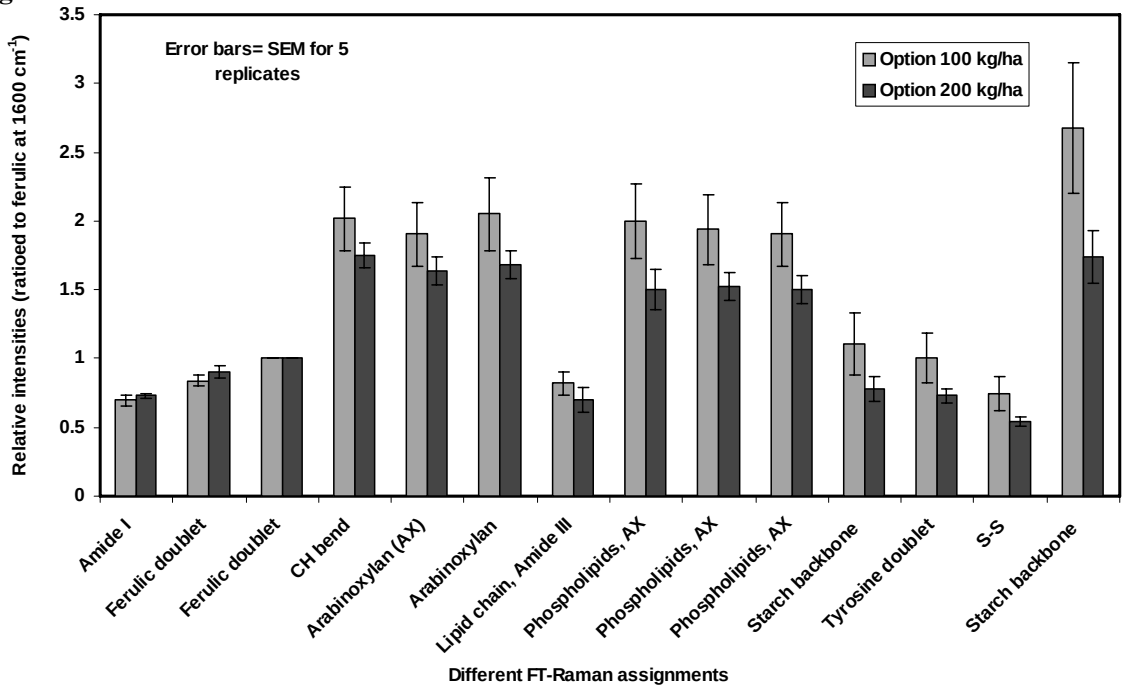
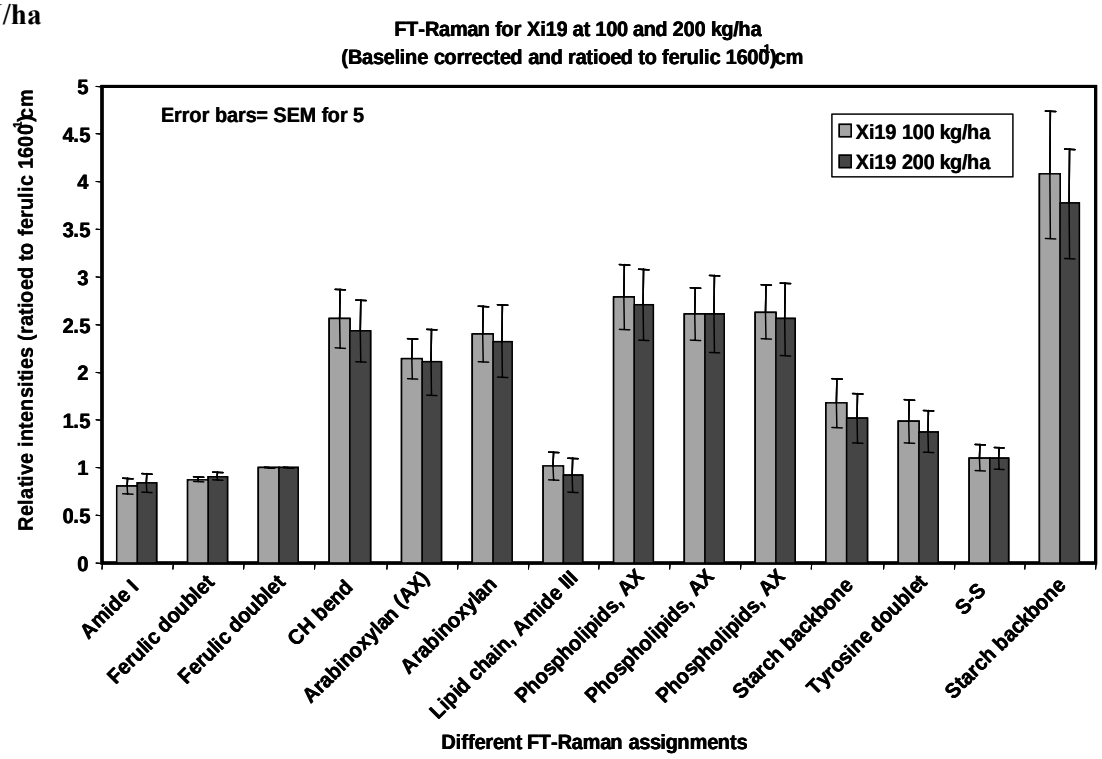


Figure 28. FT-Raman data for Xi19 flour from wheat grown with fertiliser rates of 100kg N/ha and 200kg N/ha



3.5.4 Conclusions

The objective of this part of the project was to use specialised spectroscopic techniques to investigate the impact of early N fertiliser application on the secondary structure of gluten proteins within developing and mature grain samples. The following conclusions were made:

- Lack of sample homogeneity within gel protein, protein body and thin section preparations made from developing grains resulted in poor reproducibility of the FT-IR spectral data.

- FT-IR could not detect changes in the secondary structure of gluten proteins of immature grains in response to N fertiliser application. However, principal components analysis of FT-IR spectra revealed differences in protein secondary structure based on variety.

- FT-IR spectra taken from wheat flour samples were more reproducible than those from immature material.

- FT-Raman provided information on components other than proteins. These included starch, lipids and cell wall polysaccharides. In the case of Option, there was an indication that phospholipid, arabinoxylan, starch, tyrosine and S-S contents may decrease with increasing N fertiliser input.

4. GENERAL DISCUSSION

The application of late N fertiliser in the form of foliar urea can increase protein content of wheat. The cost of the extra n fertiliser is paid for by the premium received for attaining the 13% protein content (Dumas method, dry matter basis). It has been shown that 40kg/ha extra N, applied as foliar urea at the milky-ripe development stage, can increase protein content by 1% on average (Project Report No. 109). However, it has been estimated that on average, less than 40% of the N applied as foliar urea is actually recovered in the grain (Dampney and Salmon, 1990). The remaining N is leached from the soil or lost to the atmosphere. It is known that for some wheat crops, the protein target at harvest may be reached without the benefit of additional foliar urea application (Sylvester-Bradley, 1990).

At present, growers have no ready means of assessing the N status of the crops, to help them target fertiliser decisions for boosting grain protein. A number of diagnostic or predictive systems for late N requirements for wheat crops, such as the chlorophyll meter, have been put forward. These systems have not been widely taken up on the grounds of their impracticality, poor reliability/accuracy and ability to provide only limited information. A commercial cereal tissue testing service has been operating in Australia since 1985, whereby the N content in dried rice shoots is determined by laboratory-based NIR measurements and appropriate fertiliser recommendations are sent out to the farmers. This tissue testing service was extended to other crops, including wheat. However, this service has not been widely adopted by the australian wheat industry. The reasons include the cost of the service, and the fact that excessive use of N fertilisers is rarely a problem in Australia (Batten *et al*., 2000). Such a service does not operate in the UK and appropriate NIR calibrations have not been developed for use in this country.

A previous HGCA-funded project (Bhandari, 2000) demonstrated the feasibility of predicting or forecasting the protein content of harvest grain from NIR analysis of immature grains sampled at around GS 75. See Figure A10 in Annex III. The immature ear samples were freeze-dried, de-husked and ground before being packed into cells and scanned using a laboratory-based NIR instrument. Although these results demonstrated the potential of this approach, the described protocol was clearly too slow and impractical for handling large number of samples. NIR calibrations were developed using a small data set containing only three breadmaking varieties in this earlier study. Therefore, this present study was undertaken to build upon the highly promising findings by making rapid measurements on materials with minimal transformation and to extend the range of samples. This study used the Kjeldahl method for protein determination instead of Dumas method owing to technical difficulties encountered with the designated Leco instrument during the first two years of the project. Although these issues were subsequently resolved, it was decided that the Kjeldahl method would be used for measuring N for the remaining trial years for consistency.

The dried material generally gave better NIR calibration performances for protein compared to the fresh material. However, microwave drying protocol for the immature wheat material proved to be highly labour–intensive, rather slow and required strict adherence to the prescribed method in order to avoid introducing

additional sample variability. The fresh samples of ears and whole plant had moisture contents in the region of approximately 55-80%. Despite the high moisture levels, NIR calibrations developed from the fresh or undried samples were found to be sufficiently accurate to allow them to be useful as a screening tool for assessing the status of the crop. This means that microwave drying step is not an essential requirement for determining the protein content of immature wheat by the Matrix-I.

The main objective of the project was met as we have developed a rapid NIR method for measuring protein and moisture in fresh samples of growing wheat, with minimal processing. It was found that the ears generally give better NIR calibration performance than whole plant material. This is not unexpected, as chopped ears represent a more homogeneous type of sample for the NIR measuring cell compared to chopped whole plant. Random sub-sampling from the same bulk mixture would inevitably result in a set of sub-samples with varying proportions of leaf, stem and ear segments. While it was possible to measure protein and moisture in the actual immature samples by NIR, it was more difficult to forecast the grain protein at final harvest from these data alone. This was not surprising, given that the nature of the underlying relationship between the protein in immature wheat material and harvest grain in the reference dataset, as measured using a 'wet' chemistry-based method. The squared correlation coefficient (r^2) for linear regression was 0.6 with a standard error of 1.25 for this relationship in the reference dataset. The performance of a calibration seeking to forecast harvest grain protein using NIR spectra collected from immature ear samples (fresh) was characterised with an R^2 of 0.5 with a SECV of 1.42 (standard error of cross validation), which is comparable to the analysis made for the reference dataset. In both cases, the relationship between harvest protein and that measured either directly or predicted by NIR analysis of immature material was described by a simple model, i.e. no additional factors are included.

The grain N% at harvest would have been influenced by seasonal factors, growing location, residual soil mineral N and yield. It is recognised that high yield potential tends to dilute grain N% and so some prediction of yield would also be required, in order to accurately forecast grain N%. Environmental and agronomic factors can influence yield and N uptake (Triboi and Triboi-Blondel, 2001; Jamieson *et al*., 2001). Therefore, estimates of yield potential and N uptake are important considerations when trying to predict the final grain protein. As analytical capabilities of the Matrix-I could be utilized to provide some of this information, therefore, it would be possible to develop an improved modelling system for forecasting final grain N% at harvest based the on Matrix-I, as discussed below.

A robust and generally applicable crop model could be developed based on estimates of biomass and N uptake between anthesis and milky-ripe stages and biomass accumulation during the grain filling period. The studies reported by ADAS, Boxworth showed that the crude protein in ears appeared fairly reliable between GS 61 and 79, as determined by the magnitude of deviations between observed and predicted (Annex II). These findings suggested that fundamental spectral properties of the immature plants do not change during this period. Therefore, updated matrix-I calibrations for moisture could be utilised to estimate biomass

accumulation. This would be derived from NIR moisture values and measured fresh weight per unit area. Updated matrix-I calibrations for protein would provide the estimate for N% in the plant. A realistic estimate of N uptake (kg/ha) can be made from the N% and biomass accumulation between GS 61 and GS 73. By using benchmarks for estimates of n uptake at a specific growth stage and the N harvest index (grain N content as a proportion of the whole plant content), it would be possible to forecast final grain N at harvest.

Such an approach, as outlined above, has not been possible before, using low cost, rapid sensing techniques. The Matrix-I NIR system now provides the opportunity to develop an effective tool for predicting harvest grain protein. The ability of the Matrix-I to measure protein and moisture in undried immature samples affords a major advantage over previous NIR-based methods for assessing plant tissue samples. The use of the Matrix-I will save time and labour, as well as reduce costs in the long term. While individual farmers are unlikely to meet the capital outlay for the Matrix-I instrument, it is envisaged that instruments will be acquired and operated by farming groups or co-operatives and by farming advisory organisations.

This study also investigated the application of the Matrix-I for an assessment of protein quality. The results of the NIR calibrations developed for size exclusion HPLC-resolved protein fractions were better than anticipated. Although the calibrations were derived from a limited dataset, derived from 2 trial years, they demonstrated that NIR spectral data obtained from developing wheat material could be related to the quality of harvest grain. These preliminary findings merit further investigation, particularly as there are relatively few published reports on spectroscopic analyses of gluten proteins within developing grains (Wright *et al.*, 2000). The FT-IR and FT-Raman spectroscopic techniques were undertaken to provide a better scientific understanding of how gluten proteins accumulate during grain development, and to determine the impact of genetic factors and growing conditions on protein conformation which may be related to end-use quality. FT-IR and FT-Raman methods are more sensitive than FT-NIR in providing molecular details of protein secondary structure and other components such as lipids and starches. In the present study, FT-IR analysis of thin section and protein body preparations of immature material proved to be inconclusive due to the heterogeneity of the samples. However, there was some evidence of varietal differences in the PCA of FT-IR spectral data taken from gel protein preparations of immature ear sample, indicative of variation in the conformational features of gluten proteins. Flour samples from mature grain gave reproducible spectral data, but there was little or no evidence of environmental effects in these scans.

Raman spectroscopy is particularly suitable for studying polar groups within biological systems that contain significant amounts of water, and spectra can easily be obtained for opaque samples (Ma *et al.*, 2002). The FT-Raman study performed by UEA in this project was limited to flour samples taken from harvest grain samples. Evidence of variations in starch, phospholipid, disulphide and arabinoxylan contents in response to N fertiliser was seen in some of the mature samples.

In general, growers have responded quite positively to the publicity highlighting the progress made in this project. Demonstration of the Matrix-I at various Open Day events such as those at Campden and at ADAS Boxworth, Cereals 2005 and the HGCA R&D conference in 2005 generated much interest from wheat growers.

5. REFERENCES

Batten, G.D., Blakeney, A.B., Ciavella, S. and McGrath, V.B. (2000). NIR helps to raise crop yield and grain quality. *NIR news*, **11**, 7-9.

Belton, P.S. (2005). New approaches to study the molecular basis of the mechanical properties of gluten. *Journal of Cereal Science*, **41**, 203-211.

Bhandari, D.G. (2000). The early prediction of breadmaking quality of grain and its improvement through targeted late application of nitrogen fertilisers. Project Report No.219, Home-Grown Cereals Authority, London.

Blanch, E.W., Kasarda, D.D., Hecht, L., Nielsen, K.and Barron, L.D. (2003). New insight into the solution structures of wheat gluten protein from Raman optical activity. *Biochemistry*, **42**, 5665-5673.

Büchmann, N.B., Josefsson, H. and Cowe, I.A. (2001). Performance of European artificial neural network (ANN) calibrations for moisture and protein in cereals using the Danish near-infrared transmission (NIT) network. *Cereal Chemistry*, **78**, 572-577.

CCFRA Guideline No.3. Second edition –April 1997, revised 1999. Manual of methods for wheat and flour testing. Ed. S. Salmon. CCFRA.

Dampney, P.M.R. and Salmon, S. (1990). The effect of rate and timing of late N applications to breadmaking wheats as ammonium nitrate or foliar urea-N, and the effect of foliar sulphur application. I. Effect of yield, grain quality and recovery of nitrogen in grain. *Aspects of Applied Biology*, **25**, 229-241.

Dampney, P.M.R., A. Edwards and C.J. Dyer, (2006) Managing late nitrogen applications to new Group 1 and 2 wheat varieties. Project Report No. 400, Home-Grown Cereals Authority, London.

Davy, A., Sørenson, M.B., Svendson, I., Cameron-Mills, V. and Simpson, J. (2000). Prediction of protein cleavage sites by the barley cysteine endoproteases EP-A and EP-B based on the kinetics of synthetic peptide hydrolysis. *Plant Physiology*, **122**, 137-146.

Georget, D.M.R.and Belton, P.S. (2006). Effects of temperature and water content on the secondary structure of wheat gluten studied by FT-IR spectroscopy. *Biomacromolecules*, ***7***, 469-475.

Halford, N.G., Forde, J., Anderson, O.D., Greene, F.C. and Shewry, P.R. (1987). The nucleotide and deduced amino acid sequences of an HMW glutenin subunit gene from chromosome 1B of bread wheat

(*Triticum aestivum* L.), and comparison with those of genes from chromosomes 1A and 1D. *Theoretical and Applied Genetics,* **75**, 117-126.

International Association for Cereal Science and Technology (ICC) (1986). Procedure for Near Infrared (NIR) Reflectance Analysis of Ground Wheat and Milled Wheat Products. Recommendation No. 202.

Jacobsen, S., Søndergaard, I., Møller, B., Desler, T.and Munck, L. (2005). A chemometric evaluation of the underlying physical and chemical patterns that support near infrared spectroscopy of barley seeds as a tool for explorative classification of endosperm genes and gene combination. *Journal of Cereal Chemistry*, **42**, 281-299.

Jamieson, P.D., Stone, P.J. and Semenov, M.A. (2001). Towards modelling quality in wheat – from grain nitrogen concentration to protein composition. *Aspects of Applied Biology*, **64**, 111-126.

Johansson, E., Prieto-Linde, and M.L., Svensson, G. (2004). Influence of nitrogen application rate and timing on grain protein composition and gluten strength in Swedish wheat cultivars. *Journal of Plant Nutrition and Soil Science*, **167,** 345-350.

Luo, C., Branlard, G., Griffin, W.B.and McNeil, D.L. (2000). The effect of nitrogen and sulphur fertilisation and their interaction with genotype on wheat glutenins and quality parameters. *Journal of Cereal Science*, **31**, 185-194.

Ma, C.-Y.and Phillips, D.L. (2002). FT-Raman spectroscopy and its applications in cereal science. *Cereal Chemistry*, **79**, 171-177.

Millar, S.J. (2003). The development of near infrared (NIR) spectroscopy calibrations for the prediction for wheat and flour quality.

Morel, M.-H., Dehlon, P., Autran, J.C., Leygue, J.P. and Bar-L'Helgouac'h (2000). Effects of temperature, sonication time, power settings on size distribution and extractability of total wherat flour proteinsas determined by size-exclusion high-performance liquid chromatography. *Cereal Chemistry*, **77**, 685-691.

Mutla, A., Chen, A., Reddy, S.M. and Gal, S. (1999). The aspartic proteinase is expressed in *Arabidopsis thaliana* seeds and localised in the protein bodies. *Seed Science Research,* **9**, 75-84.

Park, R.S., Agnew, R.E. and Barnes, R.J. (1999). The development of near infrared reflectance spectroscopy calibrations for undried grass silage and their transfer to another instrument using multiple and single sample standardisation. *Journal of Near Infrared Spectroscopy*, **7**, 117-131.

Piot, O., Autran, J.-C. and Manfait, M. (2000). Spatial distribution of protein and phenolic constituents in wheat grain as probed by confocal Raman microspectroscopy. *Journal of Cereal Science*, **32**, 57-71.

Piot, O., Autran, J.-C. and Manfait, M. (2001). Investigation by confocal Raman microspectroscopy of the molecular factors responsible for grain cohesion in the *Triticum aestivum* bread wheat. Role of the cell walls in the starchy endosperm. *Journal of Cereal Science*, **34**, 191-205.

Ram, M.S., Dowell, F.E.and Seitz, L.M. (2003). FT-Raman spectra of unsoaked and NaOH-soaked kernels, bran, and ferulic acid. *Cereal Chemistry*, **80**, 188-192.

Ruske, R.E., Gooding and M.H., Debraszczyk, B.J. (2004). Effects of triazole and strobilurin fungicide programmes, with and without late-season nitrogen fertiliser, on the baking quality of Malacca winter wheat. *Journal of Cereal Science*, **40**, 1-8.

Seilmeier, W., Belitz, H.-D. And Wieser, H. (1991) Separation and quantitative determination of high-molecular-weight subunits of glutenin from different wheat varieties and genetic variants of the variety Sicco. *Z. Lebensm. Unters. Forsch*., **192**, 124-129.

Shewry, P.R., Tatham, A.S. and Fido, R.J. (1995) Separation of plant proteins by electrophoresis. In: *Methods in Molecular Biology - Plant Gene Transfer and Expression Protocols,* Ed H Jones. Humana Press, Totowa, **49**, pp.399-422.

Sylvester-Bradley, R. (1990). Does extra nitrogen applied to breadmaking wheat benefit the baker? *Aspects of Applied Biology*, **25**, 217-227.

Thygesen, L.G., Løkke, M.M., Micklander, E.and Engelsen, S.B. (2003). Vibrational microspectroscopy of food. Raman *vs*. FT-IR. *Trends in Food Science and Technology*, **14**, 50-57.

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

Van Velzen, E.J.J., van Duynhoven, J.P.M., Pudney, P., Weegels, P.L.and van der Maas, J.H. (2003). Factors associated with dough stickiness as sensed by attenuated total reflectance infrared spectroscopy. *Cereal Chemistry*, **80**, 378-382.

Wellner, N., Mills, E.N.C., Brownsey, G., Wilson, R.H., Brown, N., Freeman, J., Halford, N.G., Shewry, P.R. and Belton, P.S. (2005). Changes in protein secondary structure during gluten deformation studied by dynamic Fourier transform infrared spectroscopy. *Biomacromolecules,* **6**, 255-261.

Wetzel, D.L., Srivarin, P. and Finney, J.R. (2003). Revealing protein infrared spectral detail in a heterogeneous matrix dominated by starch. *Vibrational Spectroscopy***, 31**, 109-114.

Wieser, H.and Seilmeier, W. (1998). The influence of nitrogen fertilization on quantities and proportions of different protein types in wheat flour. *Journal of the Science of Food and Agriculture*, **76**, 49-55.

Williams, M., Shewry, P.R. and Harwood, J.L. (1994). The influence of the 'greenhouse effect' on wheat (*Triticum aestivum* L.) grain lipids. *Journal of Experimental Botany*, **45**, 1379-1385.

Wright, R.J., Larroque, O.R., Bekes, F., Wellner, N., Tatham, A.S. and Shewry, P.R. (2000). Analysis of gluten proteins in developing spring wheat. In *Wheat Gluten*. *Proc. 7th Intl. Workshop Gluten 2000*. Eds PR Shewry and AS Tatham. RSC Press, pp. 471-474.

6. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support and helpful advice given by the Home-Grown Cereals Authority and by Defra. We are indebted to Bruker Optics Ltd for all their help and for making a major contribution to this project. We wish to thank the various staff members of the Cereals and Milling Department of CCFRA who contributed to the work, especially Sue Salmon, Nathalie Fa, Rachel Hurley, Fiona Organ, Doug Smith and Janet Insoll. We would like to thank Greg Hooper for his excellent work and providing expert advice on microwave drying technology. We also wish to thank Rothamsted Research staff and students. Our special thanks go to Peter Dampney and Daniel Kindred, and to the various ADAS trial managers who participated. We are grateful for the valuable input made by Richard Keeping of Heygates.

ANNEXES

ANNEX I

Figure A1. Relationship between protein (% db) at harvest and in immature whole plants (♦) and in ears (□) for 2002 trial sites at Boxworth (A), Terrington (B), High Mowthorpe (C) and Rosemaund (D)

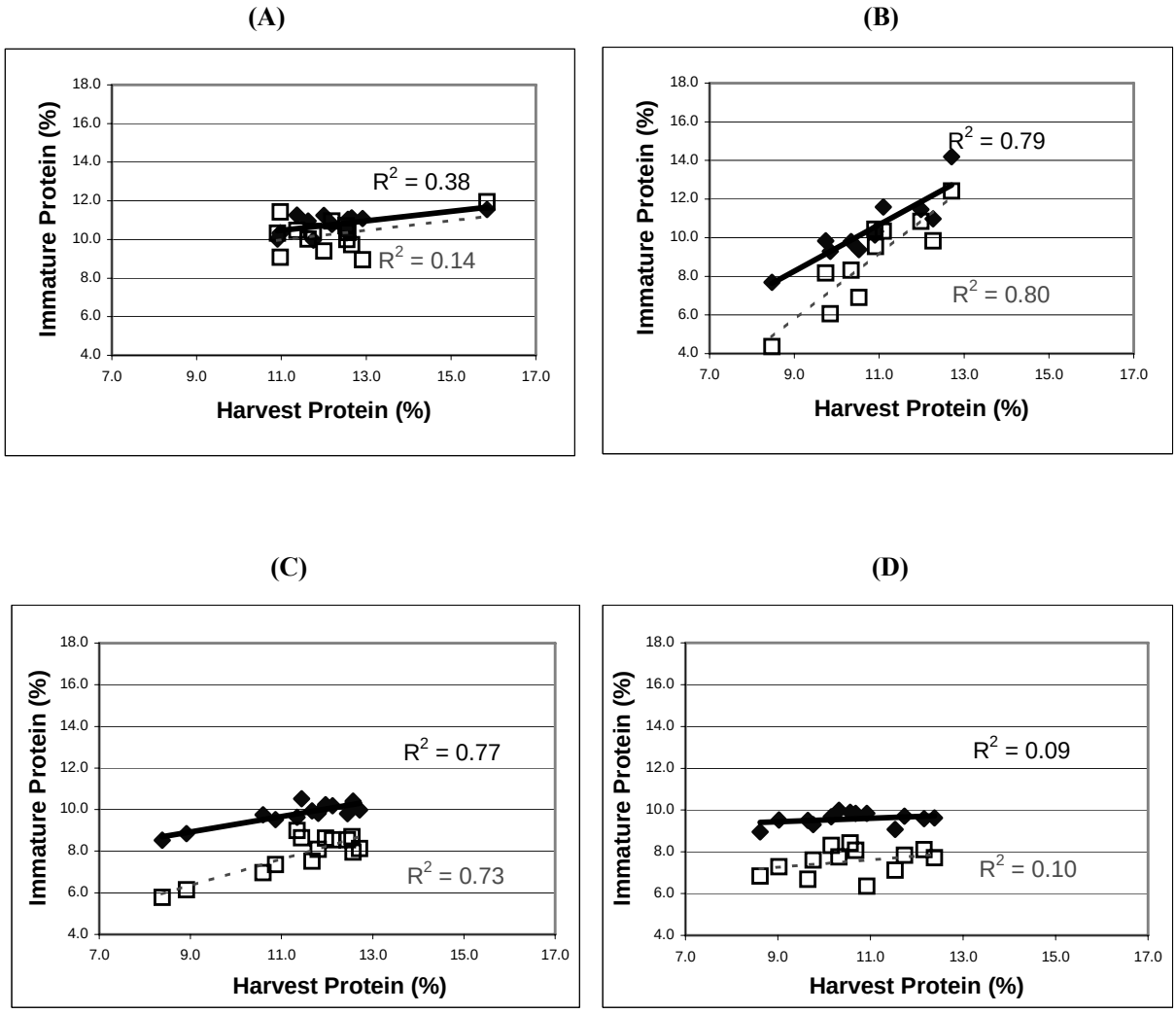
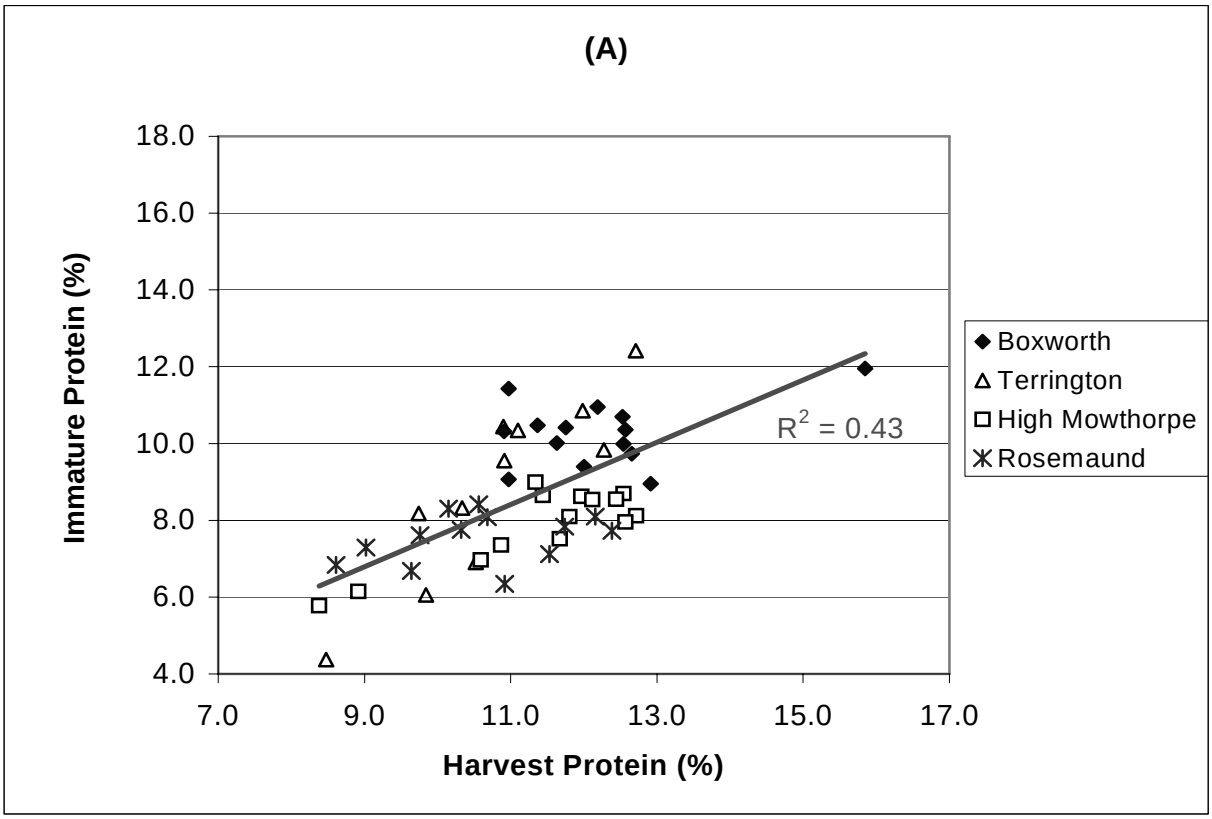


Figure A2. Relationship between protein (% db) at harvest and in immature whole plants A) and in ears B) for all trial 4 sites in 2002



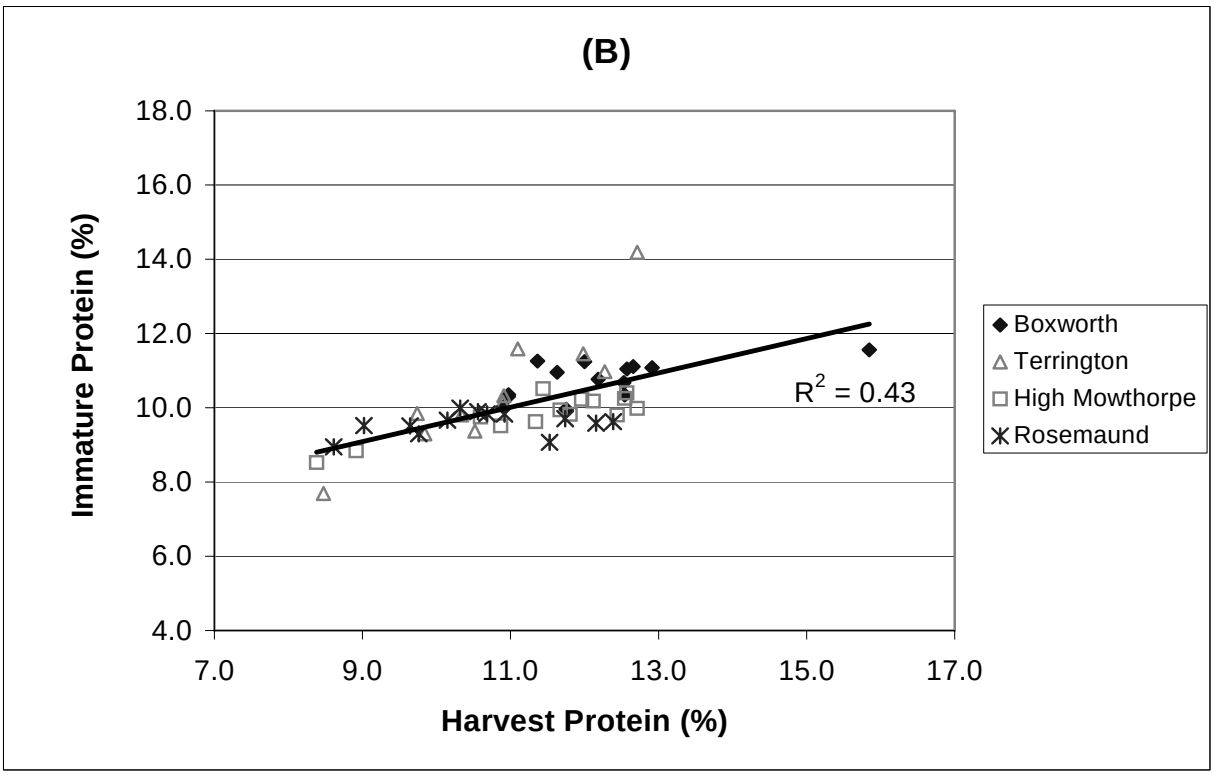
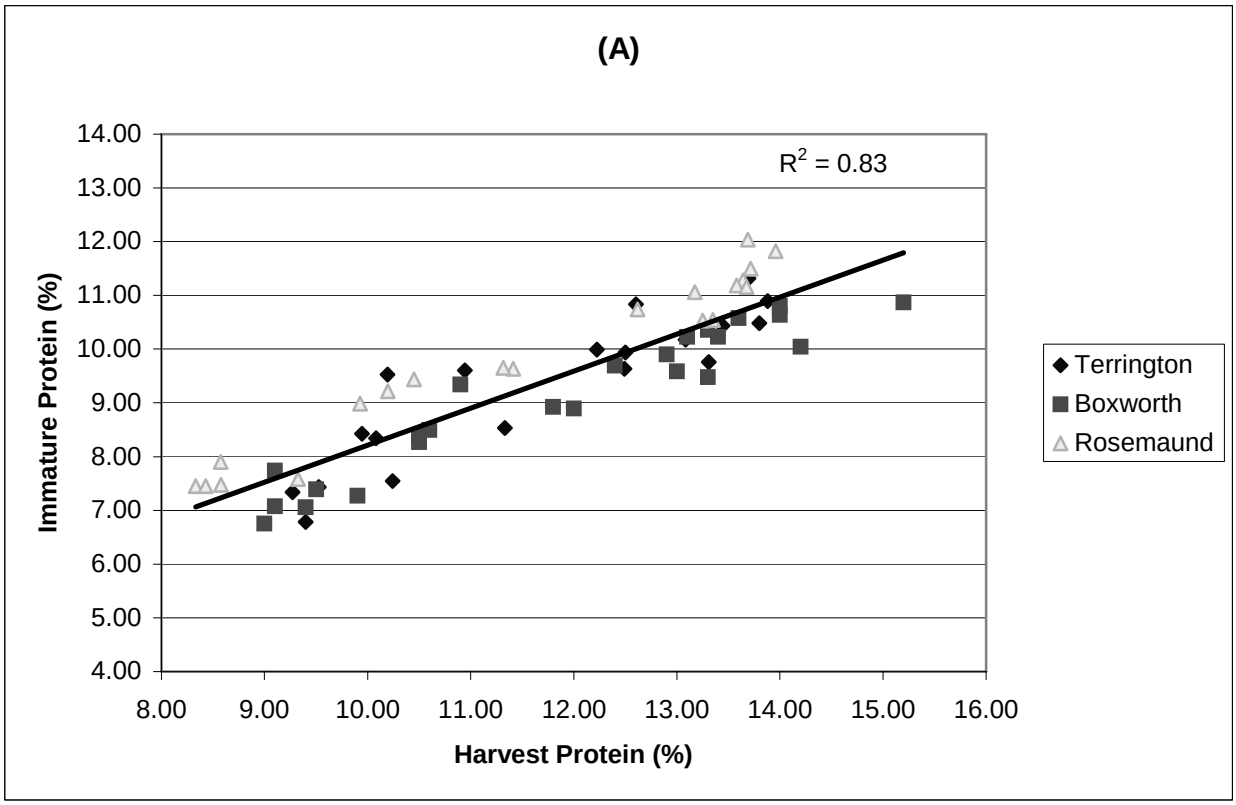


Figure A3. Relationship between protein (% db) at harvest and in immature ears A) and in whole plants B) for all trial 3 sites in 2003



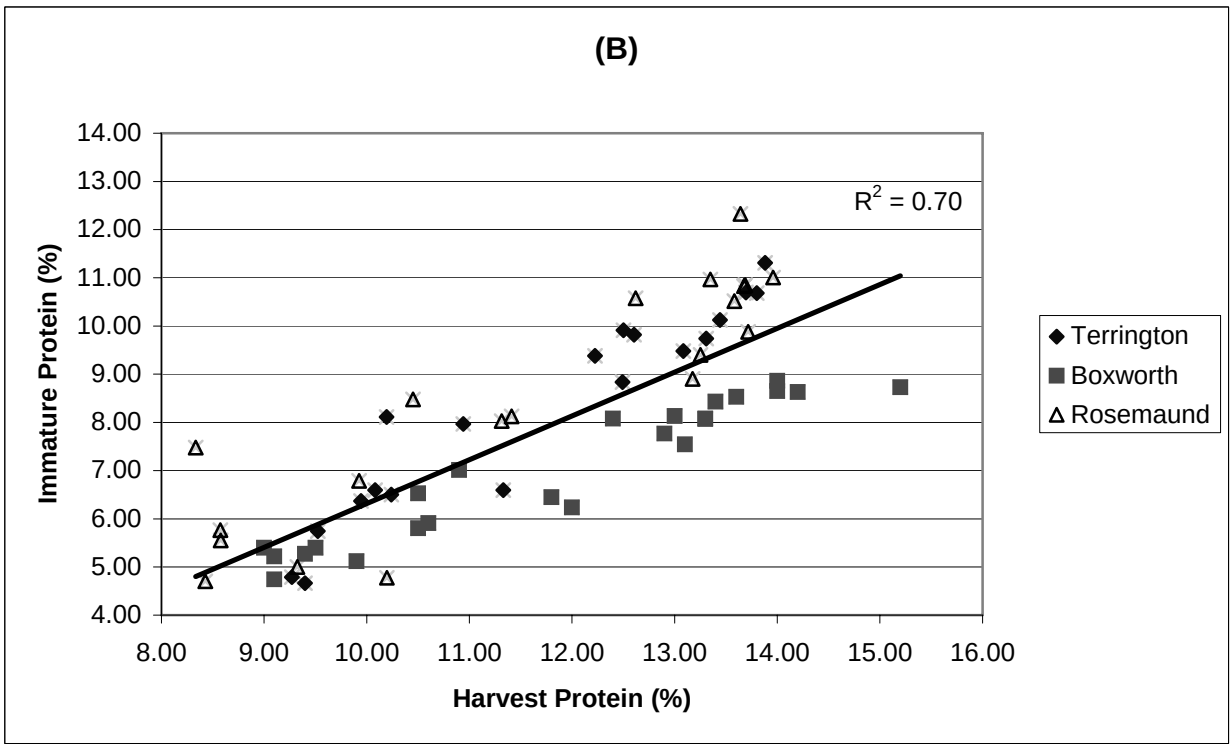


Figure A4. Relationship between protein (% db) at harvest and in immature ears for 2003 trial sites at Boxworth (A), Terrington (B) and Rosemaund (C)

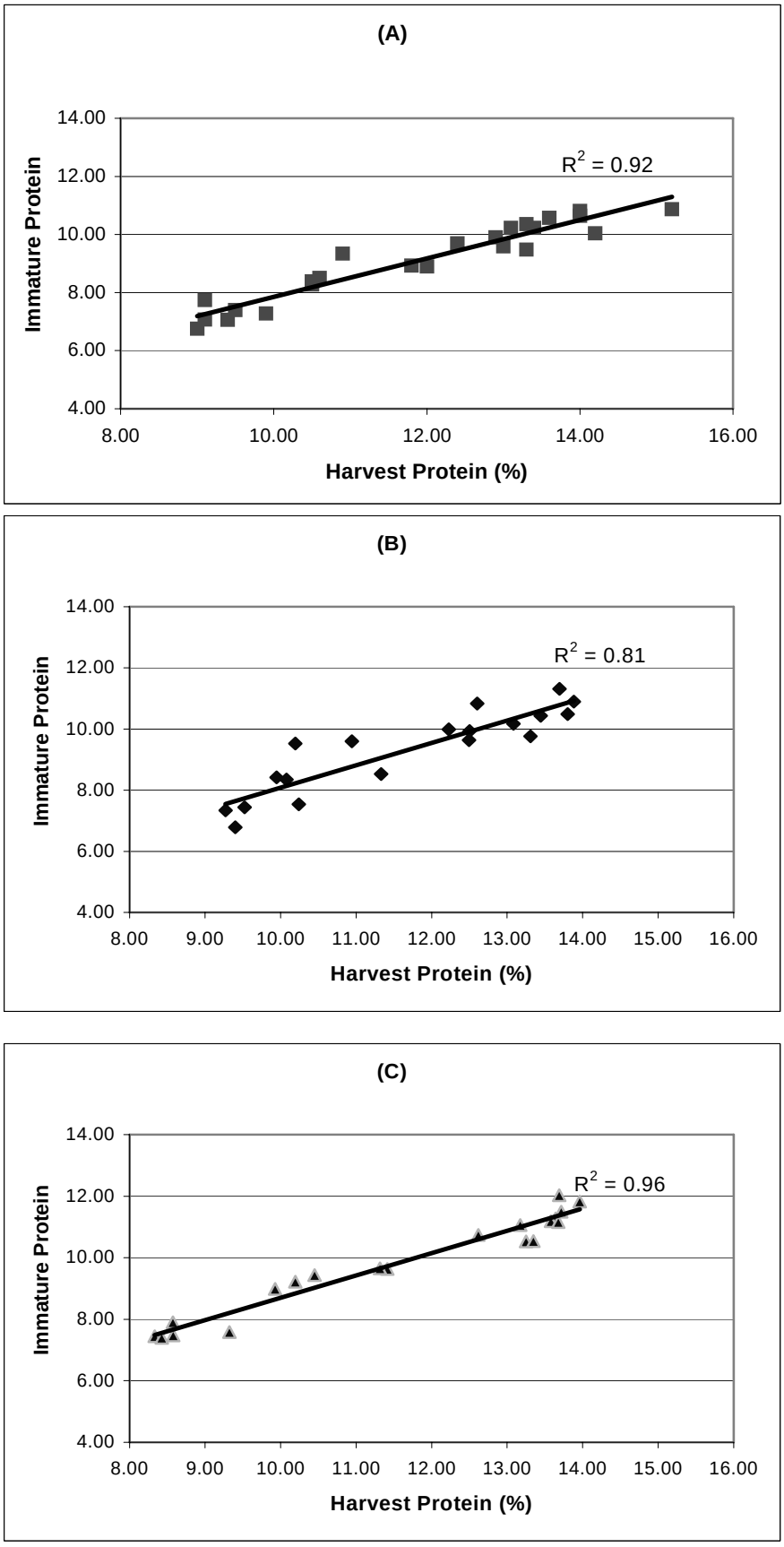
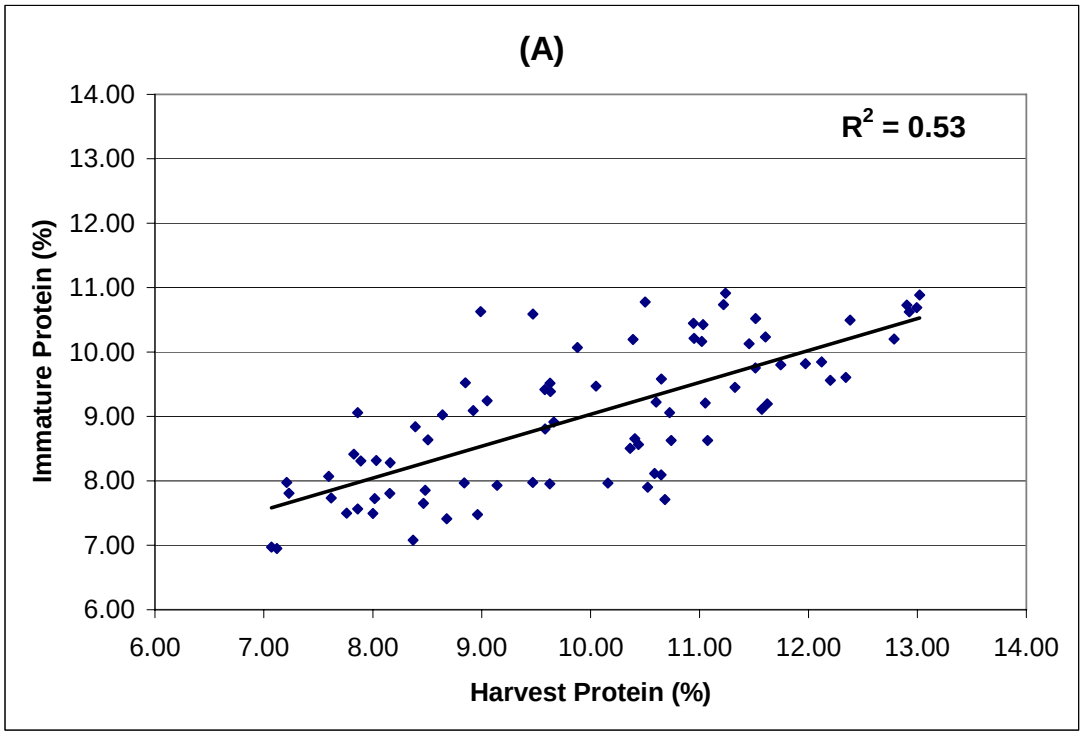


Figure A5. Relationship between protein (% db) at harvest and in immature ears A) and in whole plants B) for all trial 4 sites in 2004



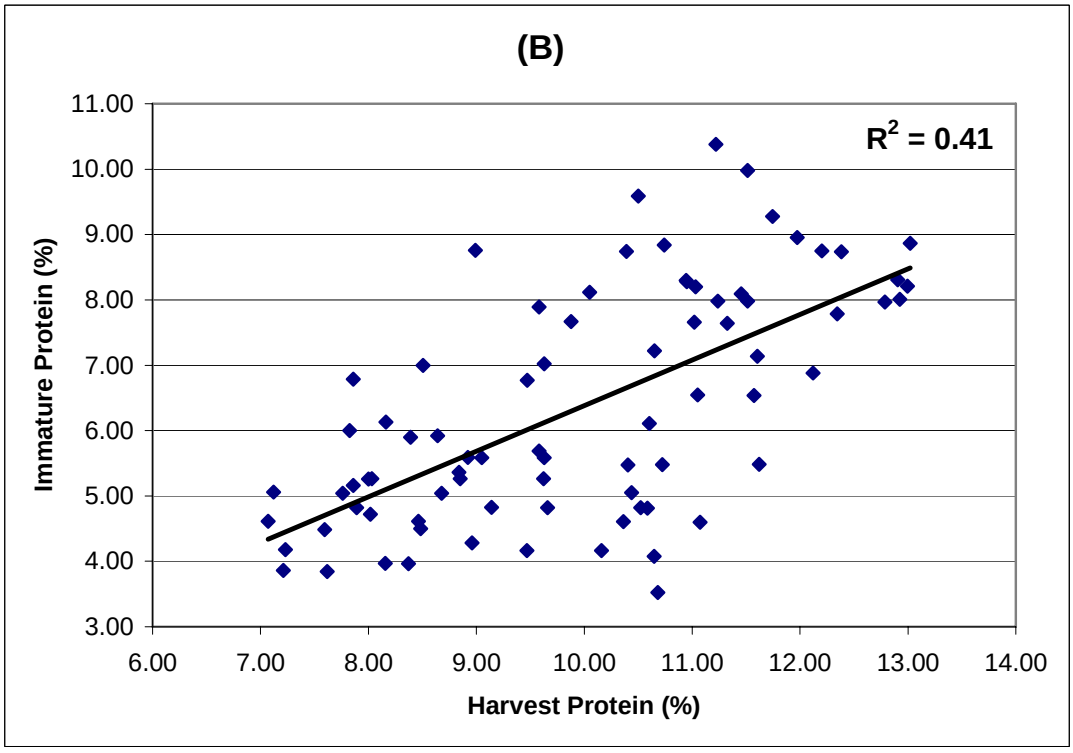


Figure A6. Relationship between protein (% db) at harvest and in immature ears (υ) and in whole plants (Δ) for the Essex trial site in 2005

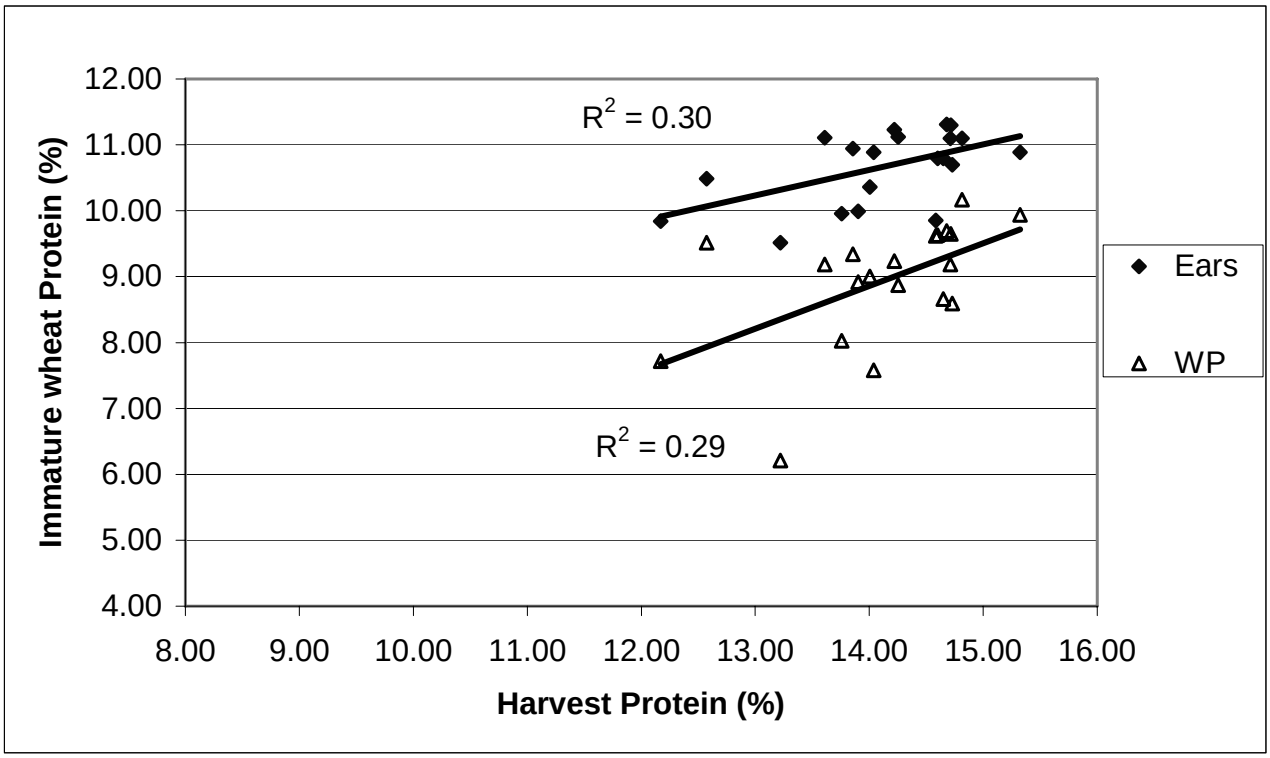
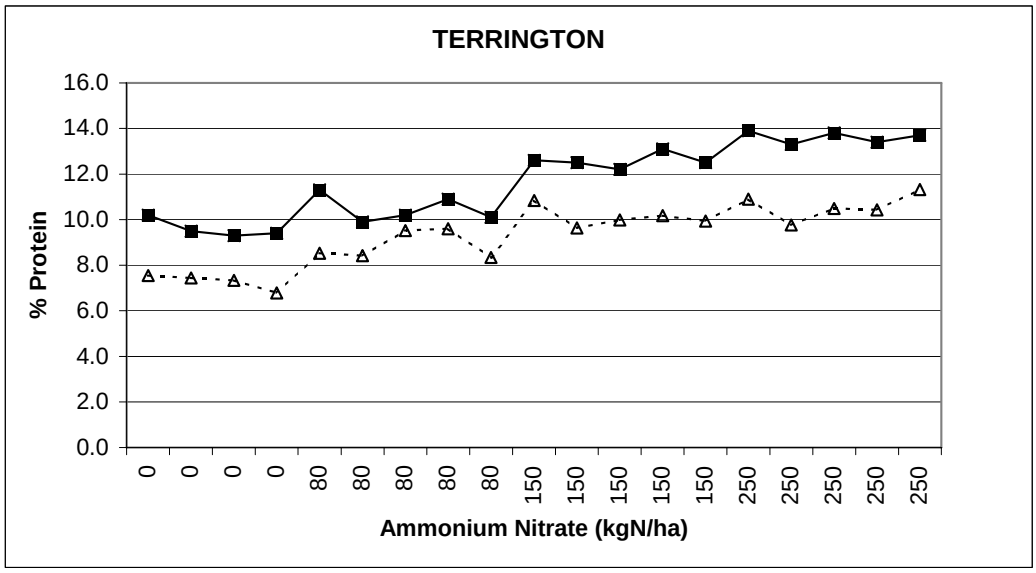
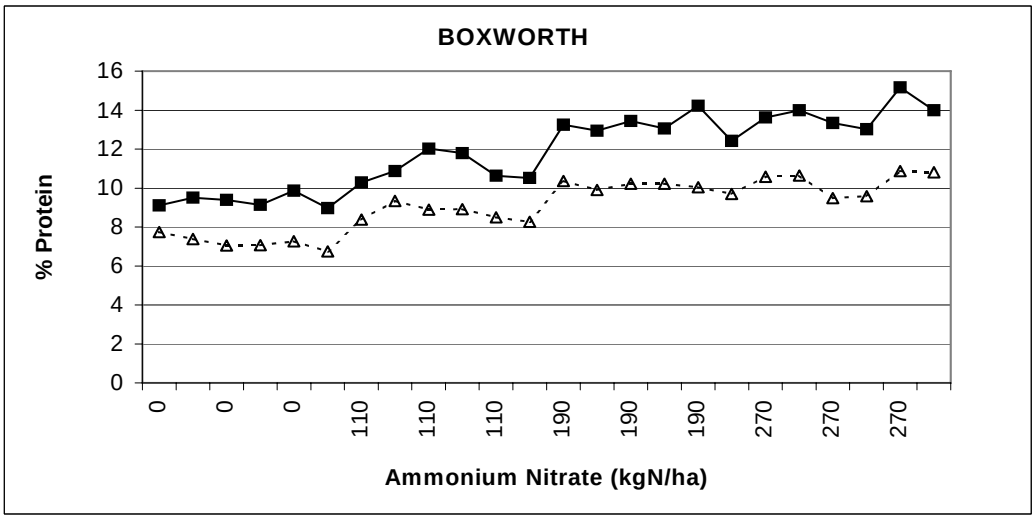


Figure A7. Effect of ammonium nitrate fertiliser treatment on the protein content (% db) of harvest grain (■) and immature ears (∆) at 3 sites in 2003 (Year 2)





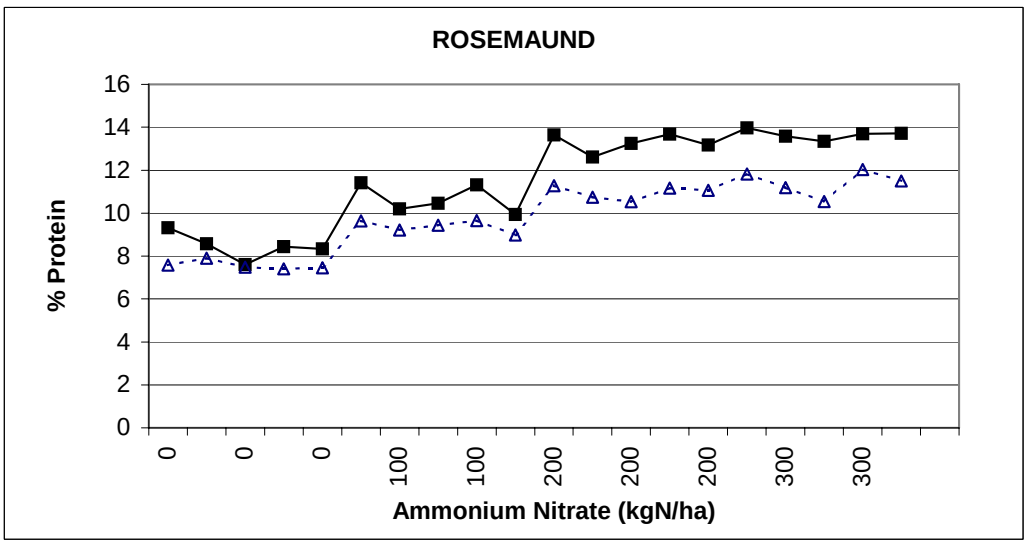


Figure A8. Effect of ammonium nitrate fertiliser treatment on the protein content (% db) of harvest grain (■) and immature whole plant (Δ) at 4 sites in 2004 (Year 3)

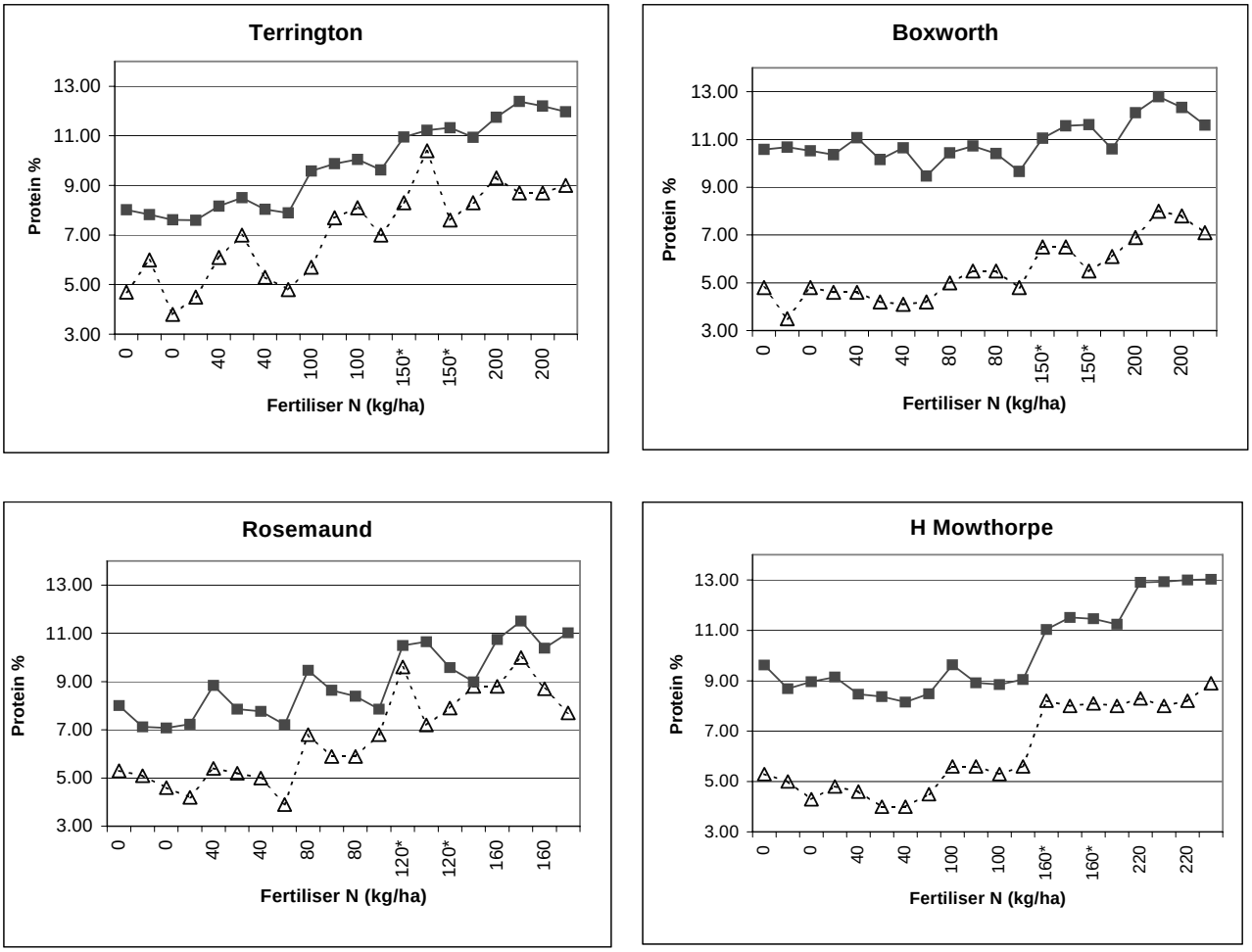
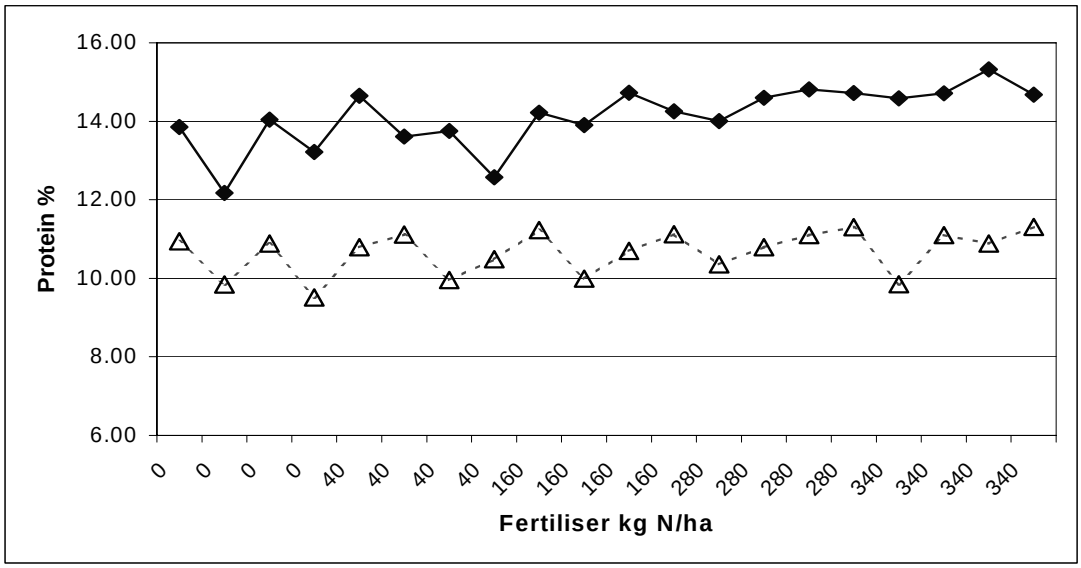


Figure A9. Effect of ammonium nitrate fertiliser treatment on the protein content (% db) of harvest grain (■) and immature ears (Δ) at the Essex trial site in 2005 (Year 4)



ANNEX II

Effects of Development Stage on Predicted and Actual Crude Protein Concentrations in Ear and Whole plant

Richard Weightman & Tamara Verhoeven, ADAS Boxworth

Background

The milky ripe stage (GS73) was the point at which all samples had been harvested during the project, in order to build the NIR calibrations for protein and moisture in immature samples. It was important to ascertain whether the NIR calibrations performed equally as well over a range of crop development stages, as they did at GS73. It would be a weakness in the calibration, if it could only be used at that specific growth stage, as there would be less flexibility for the user, for instance if a particular crop sampling date was missed. Moreover, it was recognised that there might be broader applicability if the calibrations could be used over a wider range of development stages. Therefore, a study was conducted on a crop of Xi19 at ADAS Boxworth, sampled from flowering through to early dough development, to see how actual and predicted crude protein and moisture content varied.

Materials and Methods

Data on time course sampling were collected from subsidiary plots in a field trial containing the variety Xi19. Samples were taken between 14th June and 8th July 2005 (GS 61-79, flowering to end of milky-ripe stage). Samples were taken to the laboratory, separated into two sub-samples. One sub-sample was used to make a whole plant scan using the Matrix-I. The second sub-sample was divided into ears and straw and the ears scanned. Each fraction was weighed fresh, and then dried to constant weight in a forced air-drying oven at 85^{o}C. Following drying, samples were reweighed and bagged prior to Leco N (Dumas) determination at Direct Laboratories. Crude protein was estimated from N x 5.7.

For the laboratory measurements there were three replicates from the field and the mean and SE (n=3) was estimated. From these, the upper and lower confidence limits (UCL, LCL) were estimated from the mean +/- (SE x t), where $t_{0.05, 3} = 3.18$.

For the NIR predicted values, the results of the three scans for each sample were averaged, and the three field replicates then used to calculate the SE as above. Predictions were generated by re-analysis of the stored spectra in January 2006, by Sam Millar

Results

Laboratory measurement of moisture content and Leco N determination of grain N are presented in Tables A1 and A2. It can be seen that moisture declines significantly in the ears and whole plant during these developmental stages, but less so in straw.

Table A1. Change in actual moisture content for ears, straw and whole plant in variety Xi19 from GS61-85 at Boxworth, 2005

Date	GS	Fraction	Moisture %	SE	LCL	UCL
		Ears				
14/06/2005	61	E	68.7	0.73	66.4	71.1
20/06/2005	69	E	67.3	0.30	66.4	68.3
23/06/2005	71	E	66.2	0.28	65.3	67.1
29/06/2005	73	E	66.6	0.15	66.2	67.1
05/07/2005	77	E	56.7	0.17	56.1	57.2
08/07/2005	79	E	57.5	0.33	56.5	58.6
		Straw				
14/06/2005	61	S	72.2	0.34	71.1	73.2
20/06/2005	69	S	67.9	0.77	65.4	70.3
23/06/2005	71	S	65.2	0.29	64.2	66.1
29/06/2005	73	S	*73.2	1.91	67.1	79.3
05/07/2005	77	S	66.0	0.30	65.0	67.0
08/07/2005	79	S	69.1	0.65	67.1	71.2
		Whole plant				
14/06/2005	61	WP	71.3	0.48	69.8	72.8
20/06/2005	69	WP	68.3	0.45	66.9	69.7
23/06/2005	71	WP	65.4	0.15	64.9	65.9
29/06/2005	73	WP	*70.0	0.32	68.9	71.0
05/07/2005	77	WP	62.5	0.50	61.0	64.1
08/07/2005	79	WP	63.9	0.40	62.6	65.2

*29/6/05 - Sample wet from thunderstorm on evening of 28/6/05

Protein content of straw declined markedly (Table A2), while protein content of ears and whole plant remained more constant with time. Given the limited statistical power (with n=3 observations), the data suggest that the differences over time are not significant at p=0.05. Nevertheless, they markedly show the changes occurring in straw as N is retranslocated to the ear.

Table A2. Change in actual protein content for ears, straw and whole plant in variety Xi19 from GS61-85 at Boxworth, 2005

Date	Fraction	GS	Protein (%DM)	SE	LCL	UCL
	Ears					
14/06/2005	E	61	9.3	0.39	8.1	10.6
20/06/2005	E	69	10.6	0.29	9.7	11.6
23/06/2005	E	71	10.6	0.30	9.7	11.6
29/06/2005	E	73	10.4	0.13	10.0	10.8
05/07/2005	E	77	10.0	0.16	9.5	10.6
08/07/2005	E	79	10.3	0.17	9.7	10.8
	Straw					
14/06/2005	S	61	9.1	1.09	5.7	12.6
20/06/2005	S	69	8.8	0.30	7.9	9.8
23/06/2005	S	71	9.0	0.62	7.0	11.0
29/06/2005	S	73	8.4	0.16	7.9	9.0
05/07/2005	S	77	7.5	0.69	5.3	9.7
08/07/2005	S	79	7.1	0.52	5.4	8.7
	Whole plant					
14/06/2005	WP	61	8.8	0.45	7.4	10.2
20/06/2005	WP	69	9.0	0.30	8.1	10.0
23/06/2005	WP	71	9.3	0.19	8.6	9.9
29/06/2005	WP	73	8.7	0.26	7.8	9.5
05/07/2005	WP	77	8.9	0.09	8.7	9.2
08/07/2005	WP	79	9.1	0.33	8.1	10.2

Note: Protein from 5.7 x Leco N by Direct Laboratories

NIR predictions – moisture content

NIR predictions for moisture content successfully showed the decline in moisture content for ears and whole plant (Table A3). There was a small but consistent bias, with the predicted moisture content being higher than the actual (Table A4). However, this bias may be due to the fact that the calibrations were developed from determinations of moisture made at CCFRA in previous years, whereas the actual values in Tables A3 & A4 were made in a crop drying oven which does not dry at 100^{o}C. Therefore some residual moisture may have remained in the ADAS dried samples.

NIR predictions – protein, dry matter basis

Predicted protein contents of ears and whole plant are shown in Table A5. In general, there was little consistent change in time (as seen for the actual protein values; Table A2). Note that straw was not scanned, as there were no calibrations for straw alone. Predicted and actual agreed more closely for ears than for whole plant (Table A6). From the limited data studied for ear protein, there was no evidence that the deviations were any greater at the beginning of flowering, or end of the milky ripe stage, compared to GS73.

For whole plant there was evidence for relatively large bias, i.e. an over-prediction of protein by the Matrix. Although there may be a slight bias between Leco N (actual results here) and Kjeldahl (basis of the predicted N), this would not account for deviations of this magnitude. It is likely that differences in spectral shape for 2005 samples compared with those for 2002-2004 (on which the calibrations were based at this time) accounted for the bias observed. In practice, such deviations may be easily adjusted for, using established NIR methods.

Table A3. Predicted moisture content using the Matrix-I NIR, for ears and whole plant of Xi19 grown at ADAS Boxworth in 2005

Date	GS	Fraction	Moisture (%)	SE	LCL	UCL
		Ears				
14/06/2005	61	E	71.8	0.68	69.7	74.0
20/06/2005	69	E	71.0	0.12	70.7	71.4
23/06/2005	71	E	69.5	0.19	68.9	70.1
29/06/2005	73	E	69.6	0.13	69.2	70.0
05/07/2005	77	E	60.3	0.67	58.2	62.5
08/07/2005	79	E	60.7	0.54	59.0	62.4
		Whole plant				
14/06/2005	61	WP	75.8	2.90	66.6	85.0
20/06/2005	69	WP	73.4	0.42	72.1	74.8
23/06/2005	71	WP	69.6	0.76	67.2	72.1
29/06/2005	73	WP	70.7	0.37	69.5	71.9
05/07/2005	77	WP	69.2	1.41	64.8	73.7
08/07/2005	79	WP	69.2	1.19	65.4	73.0

Table A4. Predicted and actual moisture contents (data from Tables A1 & A3) for Xi19 grown at ADAS Boxworth 2005

Moisture content %			
Predicted	**Actual**	**Predicted – Actual**	**% diff. from Actual**
	Ears		
71.8	68.7	3.1	4.5
71.0	67.3	3.7	5.5
69.5	66.2	3.3	4.9
69.6	66.6	2.9	4.4
60.3	56.7	3.7	6.5
60.7	57.5	3.2	5.6
	(Average)	(3.3)	(5.2)
	Whole plant		
75.8	71.3	4.5	6.3
73.4	68.3	5.1	7.5
69.6	65.4	4.2	6.5
70.7	70.0	0.7	1.0
69.2	62.5	6.7	10.7
69.2	63.9	5.3	8.3
	(Average)	(4.4)	(6.7)

Table A5. Predicted crude protein content using the Matrix-I NIR, for ears and whole plant of Xi19 grown at ADAS Boxworth in 2005

Date	GS	Fraction	Protein (%DM)	SE	LCL	UCL
		Ears				
14/06/2005	61	E	10.8	0.15	10.4	11.3
20/06/2005	69	E	10.4	0.15	9.9	10.9
23/06/2005	71	E	10.9	0.15	10.4	11.4
29/06/2005	73	E	9.7	0.15	9.2	10.1
05/07/2005	77	E	11.1	0.09	10.9	11.4
08/07/2005	79	E	10.2	0.18	9.6	10.7
		Whole plant				
14/06/2005	61	WP	11.4	0.77	8.9	13.8
20/06/2005	69	WP	12.7	0.37	11.5	13.9
23/06/2005	71	WP	12.4	0.69	10.2	14.6
29/06/2005	73	WP	8.9	0.01	8.9	9.0
05/07/2005	77	WP	11.8	1.09	8.4	15.3
08/07/2005	79	WP	9.4	0.63	7.3	11.4

Table A6. Predicted and actual protein contents (data from Tables A2 & A5) for Xi19 grown at ADAS Boxworth 2005

Crude protein (% DM)			
Predicted	**Actual**	**Predicted - Actual**	**% diff from actual**
	Ears		
10.8	9.3	1.5	16.3
10.4	10.6	-0.2	-2.1
10.9	10.6	0.3	2.8
9.7	10.4	-0.8	-7.3
11.1	10.0	1.1	11.1
10.2	10.3	-0.1	-1.0
	(Average)	(0.3)	(3.3)
	Whole plant		
11.4	8.8	2.6	29.0
12.7	9.0	3.7	40.6
12.4	9.3	3.1	34.0
8.9	8.7	0.3	3.0
11.8	8.9	2.9	32.3
9.4	9.1	0.2	2.6
	(Average)	(2.1)	(23.6)

Summary and Conclusions

- NIR predictions of moisture content could successfully detect the decline in ear moisture with time, as grain development progressed. Although there was a slight bias between predicted and actual, this could be an artefact of the drying method used in this study. This shows good potential for rapid prediction of DM content and, therefore biomass yield in immature material. It may also indicate that the Matrix could be used to quantify development stage during the early grain filling period.

- NIR predictions for crude protein in ears appeared fairly reliable between GS61 and 79 (as assessed by the magnitude of deviations between observed and predicted). This suggested that the fundamental spectral properties of the immature plants do not change during this period.

- In this study, crude protein content of ears was relatively stable between GS69 and 79. However, other studies have shown ear protein to decline rapidly with time during early grain filling**.**

- Rapid changes in ear protein with development stage would mean that measurements made at a 'snap shot' in time, are not likely to be useful in predicting plant N status, as the differences from week to week may be greater than the differences due to N level, unless combined with an estimate of biomass. More data is required to explore this possibility.

- The best opportunity for grain protein prediction, as discussed elsewhere, would be to combine estimates of N content and biomass to obtain an estimate of N uptake, kg/ha (which will change less with time, than will N%).

- From a physiological point of view, one would expect measurements of whole plant N to provide the best estimates of crop N uptake. However, using the current calibrations with this dataset, the predicted values for whole plant protein were *ca*. 24% higher than actual values. More recent NIR calibrations may have removed this bias.

ANNEX III

Figure A10. NIR calibration for mature wheat protein content using immature wheat spectra (taken from Project Report No. 219)