

## Validation of a curd-syneresis sensor over a range of milk composition and process parameters

M. J. Mateo,\*†<sup>1</sup> D. J. O’Callaghan,\* C. D. Everard,\* M. Castillo,‡ F. A. Payne,‡ and C. P. O’Donnell†

\*Teagasc, Moorepark Food Research Centre, Fermoy, Co. Cork, Ireland

†Biosystems Engineering, School of Agriculture, Food Science, and Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland

‡Biosystems and Agricultural Engineering, University of Kentucky, 128 C.E. Barnhart Building, Lexington 40546-0276

### ABSTRACT

An online visible–near-infrared sensor was used to monitor the course of syneresis during cheesemaking with the purpose of validating syneresis indices obtained using partial least squares, with cross-validation across a range of milk fat levels, gel firmness levels at cutting, curd cutting programs, stirring speeds, milk protein levels, and fat:protein ratio levels. Three series of trials were carried out in an 11-L cheese vat using recombined whole milk. Three factorial experimental designs were used, consisting of 1) 3 curd stirring speeds and 3 cutting programs; 2) 3 milk fat levels and 3 gel firmness levels at cutting; and 3) 2 milk protein levels and 3 fat:protein ratio levels, respectively. Milk was clotted under constant conditions in all experiments and the gel was cut according to the respective experimental design. Prediction models for production of whey and whey fat losses were developed in 2 of the experiments and validated in the other experiment. The best models gave standard error of prediction values of 6.6 g/100 g for yield of whey and 0.05 g/100 g for fat in whey, as compared with 4.4 and 0.013 g/100 g, respectively, for the calibration data sets. Robust models developed for predicting yield of whey and whey fat losses using a validation method have potential application in the cheese industry.

**Key words:** curd-syneresis sensor, validation, whey yield, whey fat

### INTRODUCTION

Experience with other bioprocesses suggests that accurate online measurement would be useful in the predictive control of cheese production. Recent developments in visible–near-infrared (NIR) optical sensor technologies can be helpful in monitoring the degree of syneresis in the cheese vat and in estimating pa-

rameters related to syneresis such as the yield of whey and curd moisture, fat, protein, and solids in whey. Previous studies have investigated the development of NIR sensors for online monitoring of the composition of milk, whey, or curd (moisture, fat, solids, and protein) over a range of experimental parameters including pH, temperature, curd stirring speed, added calcium chloride, cutting intensity, milk fat level, and gel firmness at cutting (Castillo et al., 2005; Fagan et al., 2007a,b; Everard et al., 2009; Mateo et al., 2009b). It was shown that such a sensor has potential for monitoring milk coagulation and curd syneresis simultaneously during cheese production (Castillo et al., 2005; Fagan et al., 2007a). Fagan et al. (2007b) reported that a NIR sensor response was sensitive to measured changes in syneresis kinetics as temperature was increased. Everard et al. (2009) studied the effect of milk fat level and gel firmness at cutting on the color coordinates derived from such a sensor and concluded that color parameters could be employed to predict yield of whey and fat in whey. Mateo et al. (2009b) studied the influence of curd cutting program and stirring speed on the prediction of syneresis indices and found that whey production during syneresis could be predicted using single-wavelength light backscatter (980 nm) alone.

The accuracy of the chemometric models can be assessed using different approaches including 1) cross-validation, in which a prediction of each value is made while omitting that point, or omitting a group of points including the point being predicted; 2) splitting a data set from one experiment into 2 subsets, in which one subset is used as a calibration set and the other is used as a validation set; and 3) carrying out 2 experiments, one used to develop a calibration model and the other to validate that model as in the present study. The purpose of using the third approach in this study was to know more about the potential of the online NIR sensor over a range of experimental variables and to compare the prediction accuracy of different experiments. The results obtained in this way could help in understanding how the sensor could work when installed in individual cheese plants that operate under unique conditions.

Received May 7, 2009.

Accepted July 9, 2009.

<sup>1</sup>Corresponding author: maria.mateo@teagasc.ie

To date, no previous studies in this field used this approach. The present study sought to validate the use of an online NIR sensor as a tool for predicting syneresis indices during the cheese production process. In this study, we attempted to validate predictive models by comparing data from one study against results from another study in which NIR was used to predict some important parameters during cheesemaking.

The objective of this study was to test the ability of light reflectance to predict the yield of whey and fat loss in whey over a range of experimental variables, in which 1) a calibration model was developed from one experiment and this model was validated in a subsequent experiment, 2) single-wavelength and broad-spectrum models were compared, and 3) predictive ability was compared with earlier studies involving optical reflectance.

## MATERIALS AND METHODS

### Experimental Design

Three experiments, each comprising unique experimental variables and designs, were undertaken in this study. Experiment 1 consisted of a full factorial experimental design with 3 gel cutting intensities (**CI**) and 3 curd stirring speeds (**SS**) and was undertaken in 3 replicate blocks ( $n = 27$ ). Experiment 2 consisted of a series of 27 trials that were carried out in triplicate using recombined milk with a range of milk fat levels (**FL<sub>m</sub>**; 0, 2.5, and 5.0 g/100 g) and a constant protein level (3.3 g/100 g), and the gel was cut at a range of firmness (elastic modulus, **G'**) levels (5, 35, and 65 Pa). Experiment 3 involved a series of 24 trials and was carried out using a randomized factorial design with 2 experimental factors, namely milk protein level (**PL<sub>m</sub>**) and fat:protein ratio (**F:PL<sub>m</sub>**). The PL<sub>m</sub> were 3.3 and 3.7 g/100 g, and the F:PL<sub>m</sub> were 0.3, 0.7, and 1.1. In all experiments, milk was coagulated in an 11-L cheese vat (Pierre Guerin Technologies, Mauze, France) and the gel was cut at a fixed **G'** level. Fat-level adjustment and the rheological determination of gel times were carried out according to Mateo et al. (2009a,b).

### Milk Preparation

In each of the 3 experiments, whole milk was recombined using skim milk powder, distilled water, cream (Dairygold, Cork, Ireland), and CaCl<sub>2</sub>·2H<sub>2</sub>O at 2.04 mM in a laboratory cheese vat. In experiments 1 and 2, medium-heat skim milk powder (5.39 mg of undenatured whey protein N/g of skim milk powder) was used, as described in Mateo et al. (2009a). The milk in experiment 1 had a total solids level of 12 g/100 g

and target fat and protein levels of 3.5 g/100 g and 3.3 g/100 g, respectively. However, the milk in experiment 2 had a range of FL<sub>m</sub> (0, 2.5, and 5.0 g/100 g) and a constant protein level (3.3 g/100 g).

Low-heat skim milk powder (6.25 mg of undenatured whey protein N/g of skim milk powder; Teagasc, Co. Cork, Ireland) was used in experiment 3. Also, calcium chloride was added to the milk at 2.04 mM on the day of curd production rather than on the day before curd production (i.e., the milk preparation day) as was done in the other experiments.

In each of the 3 experiments, milk was cooled to 8°C after recombination and held overnight under gentle agitation conditions (10 rpm). The formulation of ingredients in each of the 3 experiments was determined by least squares optimization using the solver tool in Excel (2002, Microsoft Corp., Redmond, WA).

### Clotting of Milk

On the day following milk preparation in each experiment, milk was heated to  $32 \pm 0.1^\circ\text{C}$  while being stirred at 22 rpm and pH was adjusted using HCl (1 M). In experiments 1 and 2, milk pH was then adjusted to 6.5 at  $32 \pm 0.1^\circ\text{C}$ . However, in experiment 3, pH was adjusted to 6.7 at  $20 \pm 0.1^\circ\text{C}$  before heating the milk to  $32 \pm 0.1^\circ\text{C}$ .

In each experiment, milk from the cheese vat was analyzed by MilkoScan (MilkoScan 605, A/S N Foss Electric, Hillerød, Denmark) to determine fat, protein, and lactose content. Rennet [Chy-Max Extra, EC 3.4.23.4, isozyme B, 600 international milk clotting units (IMCU)/mL, Chr. Hansen Ireland Ltd., Co. Cork, Ireland] was added to the milk (0.18 g of chymosin/kg of milk) in the cheese vat while being stirred constantly at 31 rpm. Stirring was stopped after 3 min and the paddles were replaced with twin cutting blades in preparation for cutting.

### Determination of Cutting Time and Gel Cutting Procedure

After stirring the rennet homogeneously into the milk, a sample of milk was removed immediately from the vat and placed in a small amplitude oscillatory rheometer, which was prewarmed to  $32 \pm 0.1^\circ\text{C}$ , to determine cutting time ( $t_{\text{cut}}$ ). Rheological analysis was carried out using a cylindrical bob and cup geometry in oscillation mode at a shear strain of 0.01 and a frequency of 1 Hz, within the linear viscoelastic region (strain <0.03) reported for rennet milk gels. The milk coagulum was cut when **G'** reached the designated level. Cutting was carried out using a twin set of cutting blades. The moment of initial gel cutting (i.e.,  $t_{\text{cut}}$ ) was taken as the

reference time ( $t$ ;  $t = 0$ ) for all subsequent syneresis-related measurements (Mateo et al., 2009a,b).

In each experiment, the cutting blades were replaced after gel cutting by the stirrers and stirring (16 rpm) commenced at  $t = 4$  min and continued over the course of syneresis.

### Sampling Procedure for the Curd–Whey Mixture

Curd and whey samples were removed using a specially designed online sampler as described in Mateo et al. (2009b). The curd–whey mixture was separated using a stainless-steel sieve and pan (Endecotts Ltd., London, UK) with a 75- $\mu\text{m}$  absolute pore size. Whey samples ( $n = 8$ ) were removed for compositional analysis at  $t = 5$  min and every 10 min thereafter up to  $t = 75$  min. The sample volume alternated between 180 and 270 mL of curd–whey mixture. The larger volume, which was sampled every 20 min from 15 to up 75 min (for a total of 4 samples), enabled analysis of fat in whey to be carried out.

### Measurements of Syneresis Indices

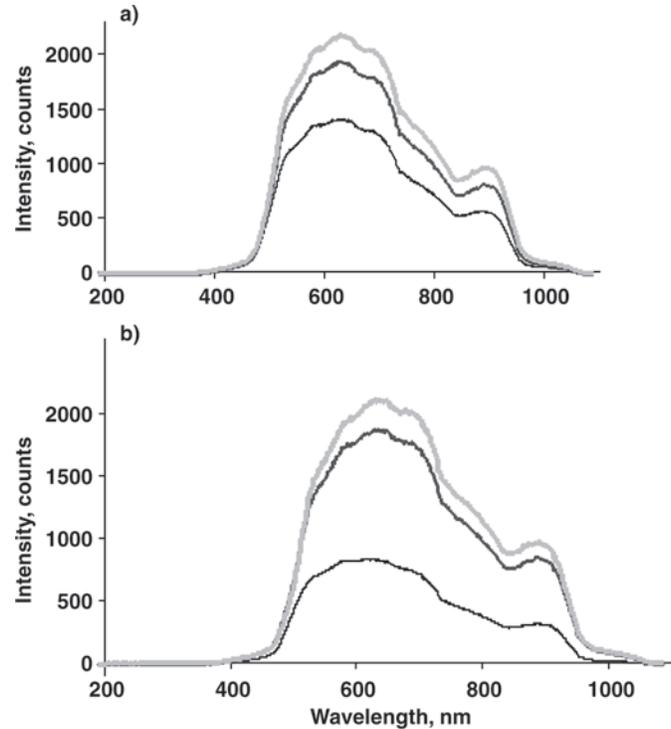
Yield of whey ( $Y_w$ ) was calculated as the percentage of whey drained immediately from each sample that was collected and whey fat ( $F_w$ ) was measured by the Rose-Gottlieb method (IDF, 1987).

### Syneresis Optical Measurements

The online NIR sensor used to carry out this study is described in Mateo et al. (2009b). The sensor used had a 6-W tungsten halogen light source (model LS1B, Ocean Optics, Inc., Dunedin, FL), which was transmitted to the mix of curd and whey through a fixed large-diameter optical fiber (diameter = 5 mm; Fiber Optics Technology, Inc., Pomfret, CT), a vertical polarizer (Edmund Optics, Inc., Barrington, NJ), and a glass window (diameter = 20 mm). Backscattered light was collected over a large area through the glass window. Reflected light was then transmitted through a second fiber (diameter = 5 mm; Fiber Optics Technology, Inc.) and a collimating lens (Edmund Optics, Inc.) that focused the scattered light onto an approximately 800- $\mu\text{m}$ -diameter fiber optic cable (Spectran Specialty Optics, Avon, CT) to the master unit of a miniature fiber optic spectrometer (HR2000CG-UV-NIR, Ocean Optics B.V., Duiven, the Netherlands), which was used as a light detector.

### Statistical Analysis

Two types of prediction models were developed involving a broad-spectrum approach and a single-wavelength



**Figure 1.** Visible–near-infrared reflectance spectra showing the effect of (a) curd-stirring speed at 3 levels [10 (thin black line), 16 (thick black line), and 22 (gray line) rpm] for one cutting intensity level (8.3 total revolutions) and (b) fat level in milk at 3 levels [0 (thin black line), 2.5 (thick black line), and 5 (gray line) g/100 g] for one gel firmness level at cutting (35 Pa). Both graphs show the untreated spectra of 3 trials during syneresis at 45 min after gel cutting.

approach, respectively. In the first approach, partial least squares (PLS) regression and cross-validation were carried out on the full spectra to develop prediction models for  $Y_w$  and  $F_w$  in all 3 experiments and to validate experiment 2 using the prediction model from experiments 1 and 3. The Unscrambler software (v9.2, Camo Process AS, Oslo, Norway) was used in the first approach. In the second approach, a single-wavelength response was extracted from the spectra at 980 nm and multiple linear regression was used to develop models for  $Y_w$  and  $F_w$  that involved either the single wavelength alone or the single wavelength combined with compositional or process parameters (i.e.,  $F_m$ ,  $G'$ ,  $F$ :PLm, PLm, SS, CI,  $t$ , and  $t_{\text{cut}}$ ). SigmaStat 3.1 software (Systat Software UK Ltd., London, UK) was used in the second approach. All models were tested for normality and constant variance.

The prediction ability of a model developed through this process is inferred from the standard error of prediction (SEP), in which  $\text{SEP}^2 = \text{mean square error (MSE)} = \text{sum of squared residuals (SSR)} / \text{degrees of freedom (df)}$ . The SEP is a measure of the accuracy of prediction models. Models were compared for pre-

diction ability across the 4 experiments, namely the 3 experiments described here and a 4th experiment described in Fagan et al. (2007b, 2008, 2009), using an *F*-test (carried out using Minitab software; Minitab 15, Minitab Ltd., Coventry, UK) for equal variance on the residuals.

## RESULTS AND DISCUSSION

### *Analysis of NIR Syneresis Spectra*

Raw spectra obtained from the NIR online sensor during syneresis at a fixed time after gel cutting in experiment 1 show an effect of curd stirring speed (Figure 1a). Similar trends were observed at all times during syneresis. It is observed that a higher stirring speed gave a higher sensor response possibly because of smaller curd particles or the larger volumes of whey at higher stirring speeds (Everard et al., 2008). This could suggest that the NIR sensor is sensitive to changes in syneresis kinetics with curd stirring speed. A similar phenomenon was observed with respect to temperature (Fagan et al., 2007b). Figure 1b shows the effect of milk fat level on the raw spectra in experiment 2. In general, there was an increase in reflectance as the fat level in milk increased, as mentioned previously.

Prediction models for  $Y_w$  and  $F_w$  were developed using the spectra (Tables 1 and 2, respectively). No treatment was carried out on the raw signals in experiments 1, 2, and 3 because the noise level was negligible, based on observation of the single-wavelength trends at 980 nm and comparison to earlier studies. Partial least squares models were developed using 2 techniques, namely 1) cross-validation, leaving out one trial at a time (Stone, 1974), and 2) jack-knifing, which consists of reducing the number of wavelengths by identifying the most significant wavelengths. This technique was applied in order to improve the robustness of the models developed in this study (Yang and Chen, 1998).

The prediction models ( $Y_w$ ,  $F_w$ ) found in this study were compared with other studies (Fagan et al., 2007b, 2008, 2009; Everard et al., 2009), which include results obtained using sensing techniques such as a broad spectrum or a single wavelength alone or in conjunction with other compositional and technology parameters or color coordinates (e.g., visible spectrum; Tables 1 and 2).

In studies in which a broad spectrum was used, the models developed for calibration in any one experiment gave a similar fit and prediction ability according to *R*, or range error ratio (**RER**). Range error ratio was calculated by dividing the range of the response variable by the SEP of the prediction model and was used to determine the practical utility of the models (Fagan

et al., 2009), using cross-validation or the jack-knifing method.

### *Prediction Models for Yield of Whey*

Comparing models for predicting  $Y_w$  using the sensing techniques mentioned, reasonably accurate predictions were obtained in experiments 1, 2, and 3 (SEP = 4.5–7.6 g/100 g; *R* = 0.83–0.92; *n* = 192–216). Analysis of residuals showed a small but significant difference in prediction accuracy between experiment 1 and experiments 2 and 3 (e.g., comparing models 1.3 and 2.3 and models 1.3 and 3.2; Table 1). This could be a result of experimental variables, different ranges of milk composition, or differences in experimental error in the individual studies.

Comparing sensing techniques (a single wavelength alone and a broad spectrum), using light backscatter alone as the means of prediction, no significant differences in prediction accuracy were found between models developed under the same experimental conditions (e.g., models 1.2 and 1.4). However, under the conditions of experiment 2, when a single-wavelength light backscatter technique was used (model 2.2), the prediction accuracy was not as high as when using a broad-spectrum approach (model 2.4). However, when the single-wavelength technique was combined with milk fat and time (model 2.1), the prediction ability was as high as that obtained using the broad-spectrum model. This suggests that a broad-spectrum approach has an advantage over a single-wavelength approach when predicting  $Y_w$  over a wider milk fat range. The use of color coordinates as a sensing technique (model 2.5) was as good as the broad-spectrum technique (model 2.3 or 2.4). Thus, a single wavelength alone is suitable for predicting yield of whey if there is no variation in milk composition. However, if milk composition varies or if there is variation in cheese technology parameters, a single-wavelength technology must incorporate other variables (e.g., milk fat level and time) or, alternatively, a broad-spectrum technology or color coordinates may be used.

In experiments 1 and 2, no differences were found in terms of the number of PLS loadings in the broad-spectrum models for yield of whey (Table 1). However, the number of loadings in experiment 3 was reduced from 4 to 3 using the jack-knifing method, giving a slight improvement in robustness although no significant improvement was detected in the other statistical parameters. In comparison with experiment 1, the calibration models 2.1 (single wavelength + milk fat + time), 2.4 (broad spectrum), 2.5 (color), and 3.1 (single wavelength + milk fat + time) for yield of whey were the most accurate in terms of *R* (0.92, 0.91, 0.92,

**Table 1.** Comparison of predictive models for volume of whey produced ( $Y_w$ ) during syneresis in various studies<sup>1</sup>

Model reference <sup>2</sup>	Parameters in model	Range of milk fat level, % wt/wt	Sensing technique	Standard error of prediction (SEP)	Correlation coefficient (R)	Range of response variable	Range error ratio (RER) <sup>3</sup>	Loadings or parameters in model, n <sup>4</sup>	Data points, n	Reference
1.1	Milk fat, curd-cutting program, light backscatter reflection ratio	3.5 <sup>5</sup>	Single wavelength (980 nm)	6.1 <sup>ab</sup>	0.91	73.5	12	3	200	Mateo et al. (2009)
1.2	Light backscatter reflection ratio	3.5 <sup>5</sup>	Single wavelength (980 nm)	6.5 <sup>a</sup>	0.89	73.5	11.3	1	208	Mateo et al. (2009)
1.3	Principal components of the spectrum (CV) <sup>6</sup>	3.5 <sup>5</sup>	Broad spectrum (189–1,100 nm)	6.4 <sup>a</sup>	0.90	73.5	11.5	5	216	Present study
1.4	Principal components of the spectrum (JK) <sup>6</sup>	3.5 <sup>5</sup>	Broad spectrum (189–1,100 nm)	6.2 <sup>a</sup>	0.90	73.5	11.8	5	216	Present study
2.1	Time after gel cutting, light backscatter reflection ratio, milk fat level	0–5	Single wavelength (980 nm)	5.2 <sup>c</sup>	0.92	60.1	11.6	3	214	Present study
2.2	Light backscatter reflection ratio	0–5	Single wavelength (980 nm)	7.6 <sup>d</sup>	0.82	55.9	7.4	1	216	Present study
2.3	Principal components of the spectrum (CV) <sup>6</sup>	0–5	Broad spectrum (189–1,100 nm)	5.3 <sup>bc</sup>	0.91	55.9	10.5	5	216	Present study
2.4	Principal components of the spectrum (JK) <sup>6</sup>	0–5	Broad spectrum (189–1,100 nm)	5.4 <sup>bc</sup>	0.91	55.9	10.3	5	216	Present study
2.5	CIE coordinates (L*, a*, b*, whiteness) <sup>7</sup>	0–5	Visible (color coordinates)	5.0 <sup>ce</sup>	0.92	63.7	12.7	4	216	Everard et al. (2009)
3.1	Light backscatter reflection ratio, time after gel cutting, milk fat	1.1–4.05	Single wavelength (980 nm)	4.45 <sup>ef</sup>	0.87	43.6	9.8	3	192	Present study
3.2	Principal components of the spectrum (CV) <sup>6</sup>	1.1–4.05	Broad spectrum (189–1,100 nm)	4.9 <sup>cf</sup>	0.83	43.6	8.9	4	192	Present study
3.3	Principal components of the spectrum (JK) <sup>6</sup>	1.1–4.05	Broad spectrum (189 nm)	4.9 <sup>cf</sup>	0.83	43.6	8.9	3	192	Present study

<sup>a–f</sup>The use of a common superscript indicates no difference in fit, as determined by a *F*-test for equal variance of the residuals.

<sup>1</sup>Each model shown on this table is a regression model applied to one experiment (i.e., experiment 1, 2, or 3).

<sup>2</sup>Format m.n, in which m denotes the experiment number and n distinguishes among models for the same experiment.

<sup>3</sup>RER = the ratio between the range of the response variable and the SEP of the prediction model.

<sup>4</sup>The broad spectrum models are partial least squares models, in which each loading corresponds to a principal component of the spectrum. There are no process or compositional parameters in those models. The other models involve the single-wavelength response with or without additional process and compositional parameters.

<sup>5</sup>SD = 0.3%, wt/wt.

<sup>6</sup>Least squares models with cross-validation (CV) or jack-knifing (JK).

<sup>7</sup>CIE = Commission Internationale de l'Éclairage. L\*, a\* and b\* refer to lightness, green–magenta, and blue–yellow in the CIE L\*, a\*, b\* chromatic value color space. Whiteness is defined by CIE<sup>7</sup> to give the best correlation between instrumental and visual assessment of whiteness.

**Table 2.** Comparison of predictive models for fat content in whey ( $F_w$ ) during syneresis in various studies<sup>1</sup>

Model reference <sup>2</sup>	Parameters in model	Range of milk fat level, %, wt/wt	Sensing technique	Standard error of prediction (SEP)	Correlation coefficient (R)	Range of response variable	Range error ratio (RER) <sup>3</sup> in model, n <sup>4</sup>	Loadings or parameters in model, n <sup>4</sup>	Data points, n	Reference
1.5	Milk fat, curd-cutting program, curd-stirring speed, time after gel cutting, light backscatter reflection ratio	3.5 <sup>5</sup>	Single wavelength (980 nm)	0.040 <sup>ab</sup>	0.77	0.28	7	5	100	Mateo et al. (2009b)
1.6	Light backscatter reflection ratio	3.5 <sup>5</sup>	Single wavelength (980 nm)	0.045 <sup>ac</sup>	0.65	0.28	6.2	1	104	Mateo et al. (2009b)
1.7	Principal components of the spectrum (CV) <sup>6</sup>	3.5 <sup>5</sup>	Broad spectrum (189–1,100 nm)	0.050 <sup>cd</sup>	0.50	0.28	5.6	2	108	Present study
1.8	Principal components of the spectrum (JK) <sup>6</sup>	3.5 <sup>5</sup>	Broad spectrum (189–1,100 nm)	0.050 <sup>ce</sup>	0.50	0.28	5.6	2	108	Present study
2.6	CIE <sup>7</sup> coordinates (L*), milk fat	0–5	Visible (color coordinates)	0.046 <sup>ae</sup>	0.94	0.50	10.9	2	72	Everard et al. (2009)
2.7	Principal components of the spectrum (CV) <sup>6</sup>	0–5	Broad spectrum (189–1,100 nm)	0.060 <sup>f</sup>	0.89	0.50	8.3	4	72	Present study
2.8	Principal components of the spectrum (JK) <sup>6</sup>	0–5	Broad spectrum (189–1,100 nm)	0.060 <sup>fh</sup>	0.89	0.50	8.2	4	72	Present study
3.4	Time after gel cutting, light backscatter reflection ratio, milk fat protein level	1.1–4.05	Single wavelength (980 nm)	0.013 <sup>i</sup>	0.99	0.36	27.7	4	96	Present study
3.5	Principal components of the spectrum (CV) <sup>6</sup>	1.1–4.05	Broad spectrum (189–1,100 nm)	0.038 <sup>aj</sup>	0.93	0.36	9.5	4	96	Present study
3.6	Principal components of the spectrum (JK) <sup>6</sup>	1.1–4.05	Broad spectrum (189–1,100 nm)	0.040 <sup>bj</sup>	0.94	0.36	9.0	2	96	Present study
4.1	Principal components of the spectrum (JK) <sup>6</sup>	3.7 <sup>5</sup>	Broad spectrum (189–1,100 nm)	0.094 <sup>k</sup>	0.91	1.41	15	5	528	Fagan et al. (2009)
4.2	Principal components of the spectrum (CV) <sup>6</sup>	3.7 <sup>5</sup>	Broad spectrum (189–1,100 nm)	0.106 <sup>lm</sup>	0.88	1.38	13	6	528	Fagan et al. (2009)
4.3	Temperature, temperature squared, milk protein, light backscatter reflection ratio	3.7 <sup>5</sup>	Single wavelength (980 nm)	0.048 <sup>adl</sup>	0.96	0.72	15	4	60 <sup>8</sup>	Fagan et al. (2008)
4.4	The time from enzyme addition to the inflexion point of the light backscatter profile (t <sub>max</sub> ), cutting time/t <sub>max</sub>	3.7 <sup>5</sup>	Single wavelength (980 nm)	0.090 <sup>m</sup>	0.59	0.37	4.1	2	20 <sup>8</sup>	Fagan et al. (2007b)

<sup>a–m</sup>The use of a common superscript indicates no difference in fit, as determined by a *F*-test for equal variance of the residuals.

<sup>1</sup>Each model shown on this table is a regression model applied to one experiment (i.e., experiment 1, 2, 3, or 4).

<sup>2</sup>Format m.n, in which m denotes the experiment number and n distinguishes among models for the same experiment.

<sup>3</sup>RER = the ratio between the range of the response variable and the SEP of the prediction model.

<sup>4</sup>The broad spectrum models are partial least squares models, in which each loading corresponds to a principal component of the spectrum. There are no process or compositional parameters in those models. The other models involve the single-wavelength response with or without additional process and compositional parameters.

<sup>5</sup>SD = 0.3%, wt/wt.

<sup>6</sup>Least squares models with cross-validation (CV) or jack-knifing (JK).

<sup>7</sup>CIE = Commission Internationale de l'Éclairage. L\* denotes lightness in the CIE L\*, a\*, b\* color space.

<sup>8</sup>Fat in whey was measured at only 85 min after gel cutting during syneresis.

**Table 3.** Statistical parameters for cross-validation models for volume of whey produced ( $Y_w$ ) and fat content in whey ( $F_w$ ), developed for experiments 1 and 3 and fitted to experiment 2

Model reference <sup>1</sup>	Validate model	Correlation coefficient (R)	Standard error of prediction (SEP)	Loadings or parameters in the model, n <sup>2</sup>	Range of response variable	Range error ratio (RER)	SD of response variable
1.3	$Y_w$	0.88	8.0	3	55.9	8.5	13.17
3.1	$Y_w$	0.87	6.6	3	63.7	9.7	13.17
3.2	$Y_w$	0.91	5.6	7	63.7	11.4	13.17
1.7	$F_w$	0.88	0.1	5	0.50	5.0	0.14
3.4	$F_w$	0.93	0.05	4	0.50	10.0	0.14
3.5	$F_w$	0.91	0.06	4	0.50	8.3	0.14

<sup>1</sup>Format m.n, in which m denotes the experiment number and n distinguishes among models for the same experiment. M.n refers to a model in Table 1 or Table 2.

<sup>2</sup>The broad-spectrum models (1.3, 3.2, 1.7, and 3.5) are partial least squares models, in which each loading corresponds to a principal component of the spectrum. There are no process or compositional parameters in those models. The other models (3.1 and 3.4) involve single-wavelength response with additional process and compositional parameters.

and 0.87, respectively) and SEP (5.2, 5.4, 5.0, and 4.5 g/100 g, respectively). This illustrates that to obtain satisfactory prediction with a single-wavelength sensor and color coordinates, it is necessary to include milk fat level as a compositional factor. However, the broad-spectrum technique does not require this additional factor because the spectrum obtained contains compositional information.

### Prediction Models for Fat in Whey

The weak correlations (0.50 to 0.77) obtained for the fat in whey models (1.5 to 1.8) in experiment 1 reflect the narrow range of the response variable investigated (Table 2). Models 2.8, 3.5, 4.1, and 4.2, which were developed using a broad-spectrum technique, were improved considerably in terms of correlation coefficient (R = 0.89, 0.93, 0.91, and 0.88, respectively) compared with model 1.7 (which was also developed using a broad-spectrum technique), although the number of PLS loadings was increased from 2 to 4 or 6 when the results from cross-validation were compared. Alternatively, when the jack-knifing method was applied, the most accurate model developed was model 3.6 (R = 0.94; SEP = 0.04 g/100 g and 2 loadings).

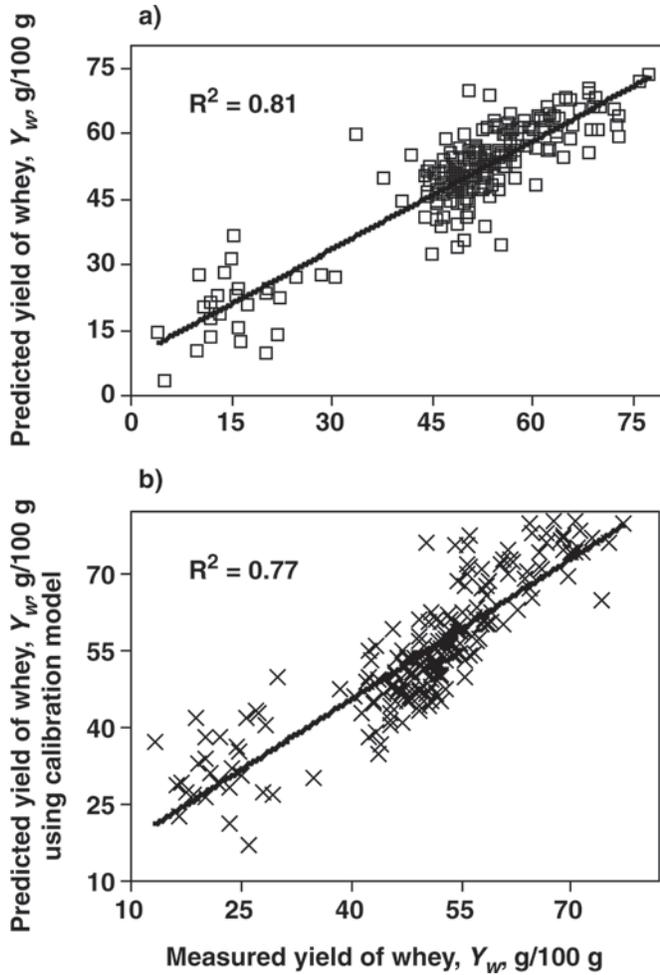
Among the models in Table 2, the lowest SEP (0.013 g/100 g whey), the highest R (0.99), and the highest RER (27.7; all found in model 3.4) were obtained over a range of milk fat:protein ratio and milk protein levels investigated in experiment 3; in this model, the sensor was used in a single-wavelength configuration in combination with designed experimental variables and time. When the milk fat level was varied in experiments 2 and 3, the single-wavelength measurement alone was not useful for predicting  $F_w$  (R < 0.1, not shown). Model 2.6, with the lightness color coordinate ( $L^*$ ) in the visible spectrum in combination with milk fat level (SEP = 0.046 g/100 g), was as good as models 1.5 (single

wavelength + milk fat + time + technology variables; 0.040 g/100 g) and 3.5 (broad spectrum; 0.038 g/100 g) in terms of SEP according to an *F*-test on the residuals. No significant differences in fit were observed among these models. Of all the models examined, the broad-spectrum technique was found to be the most robust because this technique did not require additional factors. However, a single-wavelength technique could be a more economical solution provided that variable composition and process factors are accounted for.

The extensive differences between the experimental conditions of experiment 4 and experiments 1, 2, and 3 could account for the larger SEP values in models 4.1 and 4.2, although the broad-spectrum sensing technique worked at least as well as in experiments 1, 2, and 3, in terms of R and RER. This illustrates that caution should be exercised when comparing experiments conducted in different laboratories.

There was a large range of R values among the  $F_w$  models listed, even when the SEP values were similar. The contrast between observations for SEP and for R illustrates the fact that explained variation  $R^2$  is not equivalent to prediction accuracy (SEP) when comparing experiments that differ in range of variable factors.

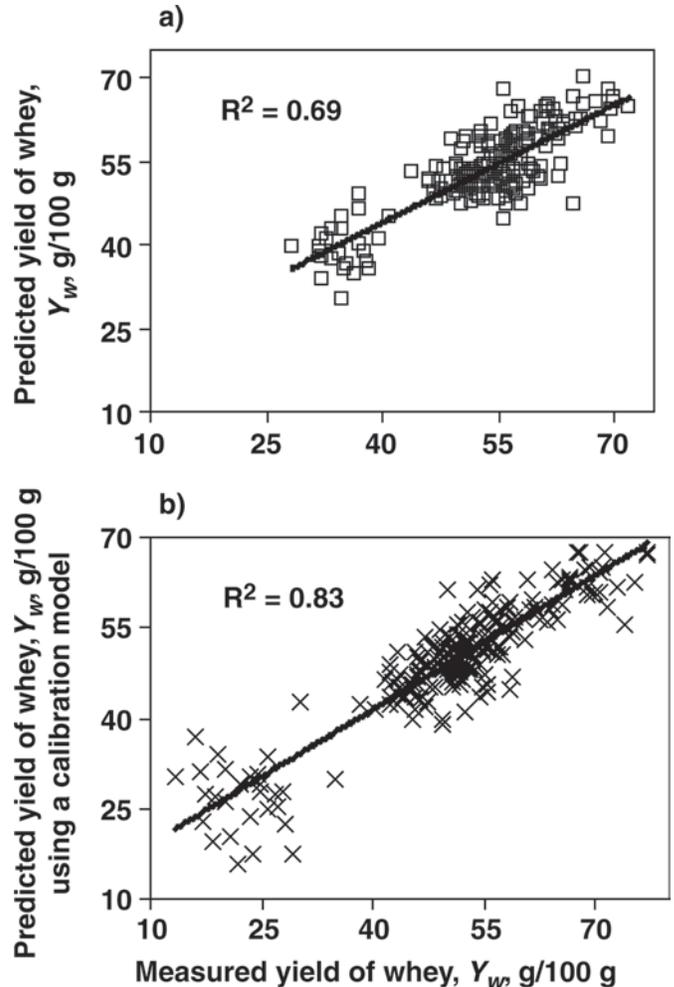
It is difficult to make a direct comparison between the use of NIR broad spectrum (models 4.1 and 4.2) and single-wavelength (models 4.3 and 4.4) models developed in experiment 4 because the single-wavelength models involve samples taken only at one time point in each trial (85 min), whereas the broad-spectrum models involve samples taken across a range of time points (5–85 min); that is, the models have different bases. Whereas the SEP for model 4.3 is less than that of model 4.1, one cannot say how they would compare on a common basis (e.g., at 85 min). This illustrates the judicious approach that is needed in making comparison between model fits, even within an experiment.



**Figure 2.** Fit of model for yield of whey,  $Y_w$ , to (a) the calibration set (experiment 1) and to (b) the validation set (experiment 2).  $Y_w$  was determined at 10-min intervals from time ( $t$ ) = 5 to 75 min after gel cutting during syneresis, giving  $n = 216$  data points in each experiment. Each experiment involves 2 clusters of data points corresponding to  $t = 5$  min and  $t > 5$  min, respectively.

**Validation of Models for Yield of Whey and Fat in Whey**

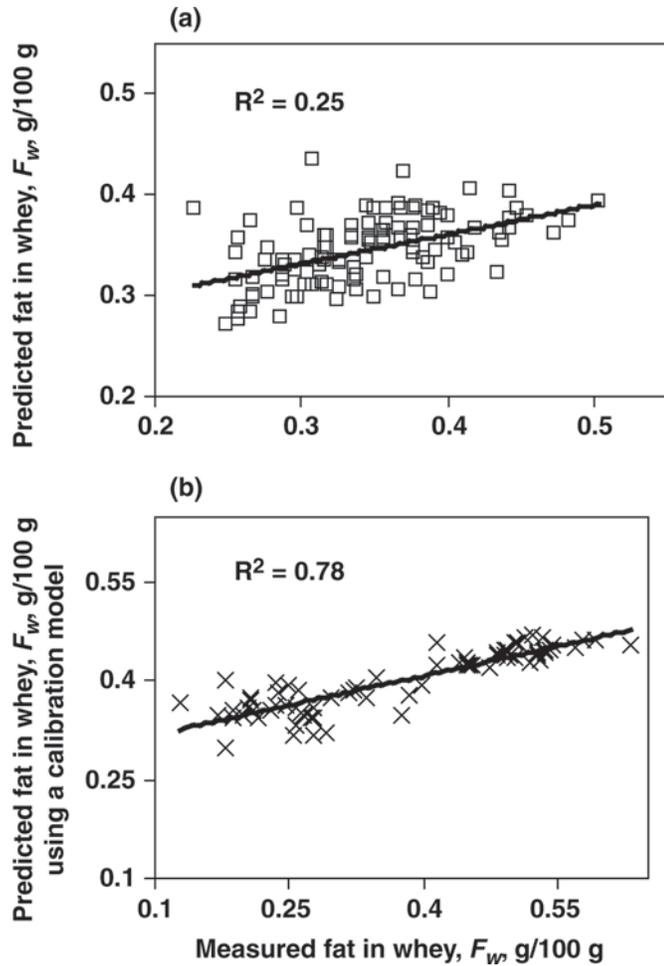
In this study, it was decided to carry out a validation test using broad-spectrum and single-wavelength (alone or in combination with compositional or technology parameters) models among the 3 experiments to determine if a calibration model obtained in one experiment can be used to validate the prediction of syneresis indices in another experiment and to confirm the usefulness of the online NIR sensor. In this case, experiments 1 and 3 were each used as a calibration set and experiment 2 was used as a validation set for each of the models in Tables 1 and 2. The models that closely fit the validation data are shown in Table 3.



**Figure 3.** Fit of model for yield of whey,  $Y_w$ , to (a) the calibration set (experiment 3) and to (b) the validation set (experiment 2).  $Y_w$  was determined at 10-min intervals from time ( $t$ ) = 5 to 75 min after gel cutting during syneresis, giving  $n = 192$  data points in experiment 3 and  $n = 216$  in experiment 2. Each experiment involves 2 clusters of data points corresponding to  $t = 5$  min and  $t > 5$  min, respectively.

Models 1.3, 3.1, and 3.2 for yield of whey, developed using the calibration set in experiments 1 and 3 (where  $R^2 = 0.81, 0.75,$  and  $0.69,$  respectively), were fitted to the validation set, experiment 2 (where  $R^2 = 0.77, 0.76,$  and  $0.83,$  respectively; Figures 2 and 3, Table 3). The data points in the 3 models fall into 2 clusters according to sampling time ( $t = 5$  min and  $t > 5$  min) as described in Mateo et al. (2009b).

In general, SEP is a more relevant measure of fit in validation than is  $R^2$  because it provides an estimate of the error in prediction. As would be expected, each model gives a slightly larger SEP with the validation set (experiment 2) than with the calibration set (experiments 1 and 3); that is, SEP increased from

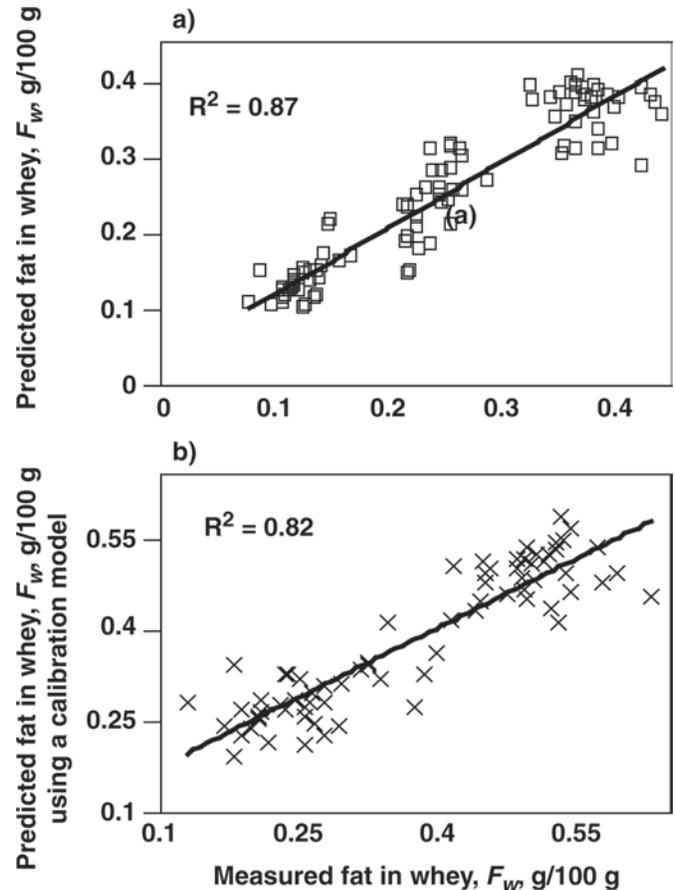


**Figure 4.** Fit of model for fat in whey,  $F_w$ , to (a) the calibration set (experiment 1) and to (b) the validation set (experiment 2).  $F_w$  was determined at 20-min intervals between time ( $t$ ) = 15 to 75 min after gel cutting during syneresis, giving  $n = 108$ .

6.4, 4.45, and 4.9 g/100 g to 8.0, 6.6, and 5.6 g/100 g for models 1.3, 3.1, and 3.2, respectively (Tables 1 and 3). Comparison of residuals (not shown) suggests no significant difference between models 3.1 and 3.2, implying that single-wavelength (combined with  $F_m$  and  $t$ ) and broad-spectrum technology were both effective at monitoring  $Y_w$  in the validation test (Table 3).

Figure 4 plots fat in whey for model 1.7 in experiment 1, the calibration set, and experiment 2, the validation set. The calibration set had a very low  $R^2$  (i.e., 0.25). Despite this, when the same model was applied to the validation set it gave a moderate  $R^2$  (i.e., 0.78). This illustrates the limitations of correlation as a measure of assessing the utility of prediction models (Workman and Mark, 2004).

However, when experiment 3 was substituted for experiment 1, using models 3.4 or 3.5 as a calibration set for validating fat in whey in experiment 2, the  $R^2$



**Figure 5.** Fit of model for fat in whey,  $F_w$ , to (a) the calibration set (experiment 3) and to (b) the validation set (experiment 2).  $F_w$  was determined at 20-min intervals between time ( $t$ ) = 15 to 75 min after gel cutting during syneresis, giving  $n = 96$  data points in experiment 3 and  $n = 108$  in experiment 2. Experiment 3 involves 3 clusters of data points corresponding to the 3 fat:protein ratios investigated.

improved significantly in the calibration set. A slight improvement was detected in the validation set (Figure 5). This could be a result of the use of different experimental variables in each experiment; for example, the milk compositional range was broadly similar in experiments 2 and 3, whereas there was only a small variation in milk fat level in experiment 1. The SEP increased from 0.050, 0.013, and 0.038 g/100 g with the calibration data to 0.10, 0.05, and 0.06 g/100 g with the validation data for models 1.7, 3.4, and 3.5, respectively (Tables 2 and 3). Models 3.4 and 3.5 were the best models for monitoring  $F_w$  in the validation test, showing that single-wavelength technology (combined with milk composition and  $t$ ) is equally as effective as broad-spectrum technology for this purpose.

In summary, the prediction of whey production and fat in whey as syneresis indices from a NIR reflectance sensor was successfully validated, based on a statistical comparison of residuals, across 3 contrasting experi-

ments in this study. Validation across experiments is not a general requirement but can be extremely useful in demonstrating the robustness of a model, especially in the context of the industrial application of a model, in which there are many in-process variations that may depart from any experimental design. The outcome of this study makes a significant contribution to that end. The limitations of this approach must of course be borne in mind. In this study, experiments 1, 2 and 3 shared some common variables (i.e., the same cheese vat, cheese technology, laboratory, and analytical methodology); however, all 3 experimental designs were very different.

### CONCLUSIONS

Comparison between calibration models suggests that an online NIR sensor was able to predict whey production and fat content in whey and showed that the use of a broad spectrum or color coordinates was superior to single wavelength when the range of milk fat content was wide, as might be expected because broad-spectrum (visible or near-infrared) sensors can measure milk fat content. Further studies will be required to determine whether combinations of a few wavelengths could provide a useful compromise between broad-spectrum, color, and single-wavelength technologies with a view to developing robust models to facilitate industrial application. In general terms, the light backscatter sensor will work in different cheese production plants but a model developed at laboratory-level or for a specific plant may need to be adapted to the needs of each unique plant, depending on the range of milk composition, cheese type, and vat design.

### ACKNOWLEDGMENTS

Funding for this research was provided under the Food Institutional Research Measure (FIRM) by the

Irish Department of Agriculture, Fisheries, and Food as part of the National Development Plan.

### REFERENCES

- Castillo, M., F. A. Payne, and A. P. Shea. 2005. Development of a combined sensor technology for monitoring coagulation and syneresis operations in cheese making. *J. Dairy Sci.* 88(Suppl. 1):142. (Abstr.)
- Everard, C. D., D. J. O'Callaghan, M. J. Mateo, C. P. O'Donnell, M. Castillo, and F. A. Payne. 2008. Effects of cutting intensity and stirring speed on syneresis and curd losses during cheese manufacture. *J. Dairy Sci.* 91:2575–2582.
- Everard, C. D., D. J. O'Callaghan, M. J. Mateo, C. P. O'Donnell, M. Castillo, and F. A. Payne. 2009. The use of colour parameters derived from an online fibre-optic sensor to monitor curd syneresis during cheesemaking. *J. Food Eng.* 94:1–6.
- Fagan, C. C., M. Castillo, D. J. O'Callaghan, F. A. Payne, and C. P. O'Donnell. 2009. Visible-near infrared spectroscopy sensor for predicting curd and whey composition during cheese processing. *Sens. Instrum. Food Qual.* 3:62–69.
- Fagan, C. C., M. Castillo, C. P. O'Donnell, D. J. O'Callaghan, and F. A. Payne. 2008. On-line prediction of cheese making indices using backscatter of near infrared light. *Int. Dairy J.* 18:120–128.
- Fagan, C. C., M. Castillo, F. A. Payne, C. P. O'Donnell, M. Leedy, and D. J. O'Callaghan. 2007a. Novel online sensor technology for continuous monitoring of milk coagulation and whey separation in cheese making. *J. Agric. Food Chem.* 55:8836–8844.
- Fagan, C. C., M. Leedy, M. Castillo, F. A. Payne, C. P. O'Donnell, and D. J. O'Callaghan. 2007b. Development of a light scatter sensor technology for on-line monitoring of milk coagulation and whey separation. *J. Food Eng.* 83:61–67.
- IDF. 1987. Milk. Determination of fat content- Röse Gottlieb gravimetric method. IDF standard 1c. International Dairy Federation, Brussels, Belgium.
- Mateo, M. J., D. J. O'Callaghan, C. D. Everard, C. C. Fagan, M. Castillo, F. A. Payne, and C. P. O'Donnell. 2009b. Influence of curd cutting programme and stirring speed on the prediction of syneresis indices in cheese-making using NIR light backscatter. *Food Sci. Technol.* 42:950–955.
- Mateo, M. J., C. D. Everard, C. C. Fagan, C. P. O'Donnell, M. Castillo, F. A. Payne, and D. J. O'Callaghan. 2009a. Effect of milk fat concentration and gel firmness on syneresis during curd stirring in cheese-making. *Int. Dairy J.* 19:264–268.
- Stone, M. 1974. Cross-validators choice and assessment of statistical prediction. *J. R. Stat. Soc., B* 39:111–133.
- Workman, J., and H. Mark. 2004. Comparison of goodness of fit statistics for linear regression. Part II. *Spectroscopy* 19:29–33.
- Yang, Z. R., and S. Chen. 1998. Robust maximum likelihood training of heteroscedastic probabilistic neural networks. *Neural Netw.* 11:739–747.