The mathematical modelling of cheese ripening

Winston L. Sweatman[1]   Steven Psaltis[2]
Steven Dargaville[3]   Alistair Fitt[4]   Tony Gibb[5]
Brodie Lawson[6]   Kaye Marion[7]

Abstract

A mathematical model is developed for the ripening of cheese. Such models may assist predicting final cheese quality using measured initial composition. The main constituent chemical reactions are described with ordinary differential equations. Numerical solutions to the model equations are found using Matlab. Unknown parameter values have been fitted using experimental data available in the literature. The results from the numerical fitting are in good agreement with the data. Statistical analysis is performed on near infrared data provided to the MISG. However, due to the inhomogeneity and limited nature of the data, not many conclusions can be drawn from the analysis. A simple

model of the potential changes in acidity of cheese is also considered. The results from this model are consistent with cheese manufacturing knowledge, in that the pH of cheddar cheese does not significantly change during ripening.

Contents

1 Introduction

The Fonterra Co-operative Group is a global company producing a variety of dairy products which are manufactured from 20 billion litres of milk per annum. The Fonterra 2013 Mathematics-in-Industry Study Group (MISG) project concerned the production, and subsequent ripening, of cheddar cheese. We desire to mathematically model changes in the gross composition of the cheese. The model will be used as a tool for predicting the final state from measured initial content of fat, protein, salt and moisture. Further, the model could help address the question of whether the process can be adjusted to improve consistent production of quality cheese to a desired specification.

Ordinary differential equations are used for constructing the mathematical model for this situation. A literature search provides details of some of the main reactions involved. Specifically, sugar (lactose), protein and fat are broken down by the starter bacteria added in the cheese-making process and the enzymes associated with them. The evolution of the fat content appears to have little effect on protein and sugar breakdown, and is modelled independently (although similar equations are used). The model has been coded in MATLAB$^{®}$ using the inbuilt non-linear optimisation routine to fit data from the literature.

The industry supplied a limited amount of data from the very early stages of ripening. Near infrared (NIR) spectroscopy is used for the measurements. Statistical analysis of this data is presented in Section 6. Relationships are found, although the data is complicated by being taken from multiple production sites with slightly varying cheddar production specifications.

The time-evolution of acidity (pH) was studied in parallel to the above model (Section 7). The pH model was consistent with information from the industry and literature that suggested little change in acidity was expected after cheese production.

The simple models presented in this paper make reasonable progress towards

producing a tool for cheese quality forecast. For further development and accurate fitting of this model, more experimental data is required over the later stages of the ripening period.

The following section gives an overview of the cheese manufacturing and quality control processes, together with a brief discussion of an existing model of cheese ripening. We then proceed to develop the model equations for the key processes in cheese ripening in Section 3, followed by a discussion of the numerical solution approach in Section 4. Section 5 gives the results obtained, and shows that we have obtained reasonable agreement with the data. In Section 6 we present a statistical analysis of the data provided by Fonterra, and in Section 7 we develop a model describing pH changes in cheese during ripening. Section 8 presents the conclusions and possible avenues for further work in this area.

2   Background

During the MISG, the Fonterra project team was given initial instruction on methods of cheese manufacture with a particular focus upon the production of a cheddar-type cheese. The cheeses are made in a batch process which takes about seven hours from the initial connection of the feedstock to the eventual production of a 20 kg block of cheese. The cheese blocks are subsequently cooled and then stored for ripening. The batch process can be run continually for about ten to eleven hours, producing several 20 kg blocks every minute.

A near infrared (NIR) spectrometer measurement is made immediately after the cheese block is produced. This measures the composition in terms of fat, moisture, salt and protein. The results of the measurement can be used to adjust the feedstock for future cheese production such as by changing the fat or protein content of the initial cheese milk used. The results of this measurement can also be used to allocate a block of cheese to various kinds of cheddar, or to cause it to be rejected. The cheese block is vacuum packed

Table 1: Acceptance limits for typical cheese products.

Property	Typical	Edam	Colby	Cheddar
Salt %	1.7	1.5–2.1	1.4–2.0	1.5–2.1
Moisture %	33	0–44	0–38	0–37
Fat %	35	0–28	33–39	33–39
Protein %	35	0–28	33–39	33–39
Lactate %	1.75			
pH	5.2	5.39–5.73	5.3–5.65	5.4–5.67

for long term storage and ripening.

The NIR measurement is repeated on core samples a few days after production (a very short period in terms of ripening time). The same quantities are measured. The information can be used to allocate the block, to reallocate the block or to cause it to be rejected.

Approximately a year later, ripened cheese blocks undergo a series of sensory tests performed by trained human assessors. These judge categories such as taste, smell, texture and crumbliness. An assessment form is completed which, along with the NIR measurements, is used by customer quality assurance to complete acceptance procedures.

Table 1 lists acceptance limits for typical cheese products.

Although there appears a lack of previous quantitative studies, a model was found for the consumption of lactose and breakdown of protein. This paper also collected data from the literature for validation [6]. The model was a good introduction to the processes involved and provided a foundation for the MISG model. However, the MISG team made further simplifications and adaptations during the project week to develop their own model describing bacteria, lactose, enzymes, protein and peptides in the cheese. In addition, the breakdown of fat has been included.

One particular aspect of the earlier study [6] was a sharp division within

the model into two time phases and fitting with different values of model parameters. The first phase lasts about two to four weeks, and is a period in which numbers of bacterial cells are undergoing substantial change. There is an initial growth as residual lactose is consumed but then a steady decline due to cell death and lysis. Enzymes are produced by the bacteria. In the second phase, bacterial activity has essentially ceased, and the ripening is dominated by chemical reactions catalysed by the enzymes. The new model seeks to work with the same consistent set of model equations, parameters and initial conditions throughout the whole ripening process. The different phases are still represented and can be identified in the numerical comparisons with data.

3   Model equations for maturing of cheese

3.1   Modelling discussion

We now proceed to develop model equations for the process of cheese ripening and maturing. Clearly the process of cheese ripening is extremely complicated involving a great many processes and reactions, and many different modelling choices could be made. One of the key modelling decisions regards which quantities to keep track of, and which to relegate as being unimportant in the final cheese product. Our choices and much of our modelling are based on the work of Kim et al. [6], though the final set of equations that we have used is substantially different.

Our main modelling hypotheses include the assumptions that, though cheese is composed of fat, protein, water, salt and lactose,

1. the salt content remains unchanged during the maturation process, and

2. the water and the fat are not primary reagents (and in particular, the fat is not a controlling parameter).
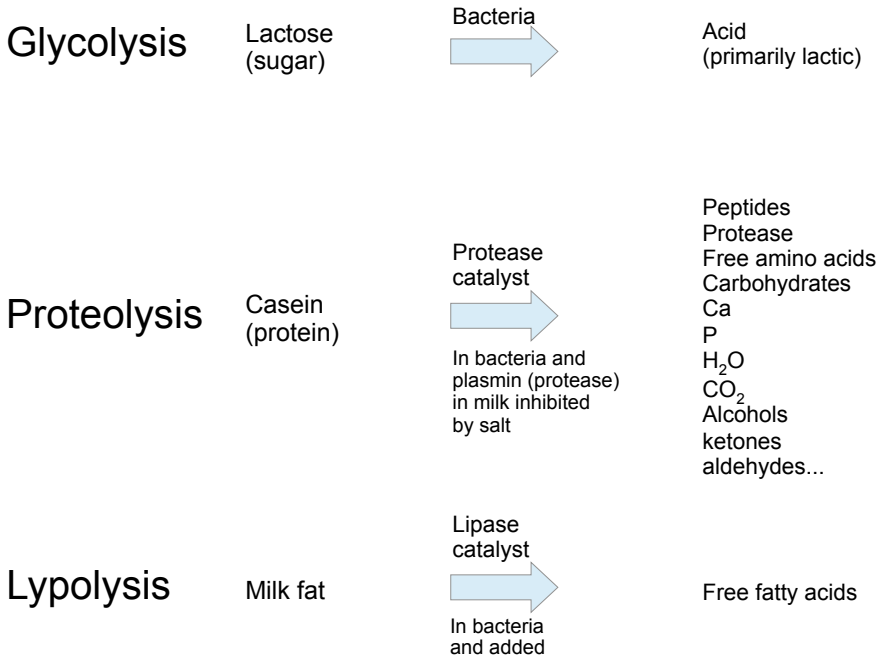
Figure 1: Three key processes by which the gross composition of cheese changes. As detailed by Marsili [9], there are multiple products in each case.

We also assume that cheese maturation and ripening involves three main processes (Figure 1): glycolysis (where sugar—in this case lactose—feeds bacteria and produces acids including lactic acid), proteolysis (where protein—in this case casein—is broken down by protease catalysts to form peptides, free amino acids, carbohydrates, alcohols, water, CO_2 and a range of other products) and lipolysis (where milk fat is broken down by catalysts to form free fatty acids). The first two of these processes are modelled in equations (1)–(8) below, and the third in equations (9)–(11).

In the brief descriptions below we explain why we believe a given quantity is important to model, and also give some details of our reasoning behind the final differential equation model that was eventually adopted. Table 2

summarises the nomenclature used in the model equations.

Bacterial cells: the bacterial cells that cause cheese ripening are key to the whole process of maturation. They are measured in units of "colony forming units" per gramme (cfu/g), thereby involving a measure of viable bacterial cells. We assume (as many others before have done) that the cells evolve according to classical Michaelis–Menten kinetics, growing via consumption of lactose. When the amount of lactose present is small the bacterial growth rate is proportional to the amount of lactose present, multiplied by a constant growth factor μ_m/K_ℓ. However, when lactose saturates, that is becomes large relative to the reaction constant K_ℓ, its amount becomes unimportant and the bacterial growth rate is determined by a constant μ_m. We also assume that the bacterial cells die with a constant known death rate k_ℓ per day.

Lactose: lactose ($C_{12}H_{22}O_{11}$) is consumed by the bacterial cells as food for growth. Since we measure the amount of lactose in units of mg/g, a constant scaling factor Y_x of dimensions cfu/mg must be employed. Lactose is also lost through conversion to lactic acid via a Michaelis–Menten reaction with saturation constant μ_{LA}.

Lactic acid: lactic acid ($C_3H_6O_3$) is a byproduct of the breakdown and consumption of lactose by the bacterial cells. It plays a key role in the souring (or fermentation) of the milk constituent of cheese. Lactic acid is assumed to be produced via a Michaelis–Menten reaction involving lactose and bacterial cells with the saturation constant μ_{LA} and an additional factor κ_1 to allow for the differences in mass between lactose consumed and lactic acid produced. The production of acid and implications for pH are discussed further in Section 7.

Table 2: Nomenclature used in model equations

Symbol	Description
A	amount of casein $(\mathrm{mg\,g^{-1}})$
α	amount of lactic acid $(\mathrm{mg\,g^{-1}})$
B	amount of dipeptides $(\mathrm{mg\,g^{-1}})$
C	amount of amino acids $(\mathrm{mg\,g^{-1}})$
E_1	amount of proteinase $(\mathrm{U\,g^{-1}})$
E_2^{o}	amount of extracellular dipeptidase $(\mathrm{U\,g^{-1}})$
E_L	amount of extracellular lipase $(\mathrm{U\,g^{-1}})$
F	amount of milk fatty acids $(\mathrm{mg\,g^{-1}})$
L	amount of lactose $(\mathrm{mg\,g^{-1}})$
T	amount of triglycerides $(\mathrm{mg\,g^{-1}})$
X	bacterial cell concentration $(\mathrm{cfu\,g^{-1}})$
t	time (days)
μ	dipeptidase reaction constant U $\mathrm{cfu^{-1}}$
μ_m	cell reaction constant $(\mathrm{day^{-1}})$
μ_L	lipase reaction constant $(\mathrm{U\,cfu^{-1}})$
μ_{LA}	lactic acid reaction constant $(\mathrm{mg\,cfu^{-1}\,day^{-1}})$
k_ℓ	cell death rate $(\mathrm{day^{-1}})$
k_1	extracellular dipeptidase destruction rate $(\mathrm{day^{-1}})$
k_2	proteinase destruction rate $(\mathrm{day^{-1}})$
k_3	lipase destruction rate $(\mathrm{day^{-1}})$
K_A	reaction constant $(\mathrm{mg\,g^{-1}})$
K_B	reaction constant $(\mathrm{mg\,g^{-1}})$
K_T	triglyceride reaction constant $(\mathrm{mg\,g^{-1}})$
K_ℓ	reaction constant $(\mathrm{mg\,g^{-1}})$
V_f	casein reaction constant $(\mathrm{mg\,U^{-1}\,day^{-1}})$
V_b	amino acids reaction constant $(\mathrm{mg\,U^{-1}\,day^{-1}})$
V_T	triglyceride reaction constant $(\mathrm{mg\,U^{-1}\,day^{-1}})$
Y_x	lactose yield constant $(\mathrm{cfu\,mg^{-1}})$
α_1	proteinase reaction constant $(\mathrm{U\,cfu^{-1}})$
ζ	scaling constant for proteolysis reactions
κ_1	constant of proportionality relating lactose and lactic acid
κ_2	constant of proportionality relating triglycerides and fatty acids

Proteinase: in order to ensure that the necessary reactions proceed, it is essential that a proteinase enzyme is produced by the bacterial cells. Tests have shown [7, e.g.] that proteinase may markedly influence the final flavour of cheese, so it is important to predict its development during the maturation process. We assume that proteinase is produced via a Michaelis–Menten type reaction involving bacterial cells and lactose, with a saturation constant $\alpha_1 \mu_m$, and also suffers a natural destruction rate with constant of proportionality k_1. Since it is an enzyme, proteinase is measured in units of U/g—enzyme units per gramme ($1\,U = 1\,\mu mol/min = 16.67$ nanokatals).

Extracellular dipeptidase: the dipeptidase enzyme seems to exist in ripening cheese in both intracellular and extracellular forms. Since it appears likely that the extracellular dipeptidase is the main accelerant in the production of amino acids, we model this component here, assuming that it is produced from bacterial cells at a rate μk_ℓ, and is destroyed at a rate with constant of proportionality k_2. Dipeptidase is also measured in standard enzyme units per gramme (U/g).

Casein: Casein (phosphoprotein) is the dominant protein found in milk, and plays a crucial role in encouraging clot formation during the cheese-making process. Since it has long been suspected that casein is responsible for some of the distinctive characteristics of matured cheese, it should be tracked during the cheese manufacturing process. We assume that it is consumed via a Michaelis–Menten type reaction involving the proteinase enzyme, with a saturation constant V_f.

Amino acids: for many cheeses the presence of tyrosine clusters is a key component in the maturity and "ripeness" of the cheese. These clusters are formed by amino acids, so it is important to track the formation of these acids. We assume that amino acids are formed via a Michaelis–Menten type reaction between dipeptide and extracellular dipeptidase, with a saturation

constant V_b. More complicated models of the cheese maturation process could include submodels for individual amino acids that are thought to be particularly important components in the taste and smell of matured cheese.

Dipeptides: some studies [13, e.g.] have indicated that dipeptides play a key role in aspects of the taste properties of matured cheese. The dipeptide content of the cheese decreases as its casein and amino acid components increase, with simple multiplicative factors ζ and ζ^{-1}, respectively. These factors account for the differences in molecular weights of these protein molecules and additional water molecules during hydrolysis reactions [6].

3.2 Working equations for cheese maturation

The above considerations lead to the following set of first-order ordinary differential equations which describe conservation of mass for what we judge are the most important quantities in the cheese maturation process:

Bacteria cells X

$$\frac{dX}{dt} = \frac{\mu_m L X}{K_\ell + L} - k_\ell X \, ; \tag{1}$$

Lactose L

$$\frac{dL}{dt} = \left(-\frac{\mu_m}{Y_x} - \mu_{LA} \right) \frac{L X}{K_\ell + L} \, ; \tag{2}$$

Lactic acid α

$$\frac{d\alpha}{dt} = \kappa_1 \frac{\mu_{LA} L X}{K_\ell + L} \, ; \tag{3}$$

Proteinase E_1

$$\frac{dE_1}{dt} = \frac{\alpha_1 \mu_m L X}{K_\ell + L} - k_1 E_1 \, ; \tag{4}$$

Extracellular dipeptidase E_2^o

$$\frac{dE_2^o}{dt} = \mu k_\ell X - k_2 E_2^o \; ; \tag{5}$$

Casein A

$$\frac{dA}{dt} = -\frac{V_f E_1 A}{K_A + A} \; ; \tag{6}$$

Amino acids C

$$\frac{dC}{dt} = \frac{V_b E_2^o B}{K_B + B} \; ; \tag{7}$$

Dipeptides B

$$\frac{dB}{dt} = -\zeta \frac{dA}{dt} - \frac{1}{\zeta} \frac{dC}{dt} \; . \tag{8}$$

3.3 Additional equations for lipolysis

It is well known that the flavour of cheese is affected by the fat content [2] and so we add equations modelling the evolution of fat content in cheese in our model. In the process of lipolysis, milk fat is broken down into free fatty acids. It seems clear that a key role is played by the lipase enzyme added during production, which is present in starter bacteria and released during lysis. It also seems reasonable to assume that though lipolysis is a process that is coupled to both glycolysis and proteolysis, there is no back-coupling and hence the lipolysis equations may be solved as an independent system once the other quantities are known. We model the extracellular lipase enzyme E_L which is produced by bacterial cells, measured in units of U/g, and subject to a standard enzymatic destruction rate with constant of proportionality k_3. Milk fat is considered in the form of triglycerides T whose conversion to fatty acids F is controlled via a Michaelis–Menten reaction catalysed by the extracellular lipase. As in equations (3) and (8) an additional factor (κ_2) has been included to allow for the differences in mass between triglycerides

consumed and fatty acids produced. The working equations for lipolysis are taken to be

Extracellular lipase E_L

$$\frac{dE_L}{dt} = \mu_L k_\ell X - k_3 E_L \, , \tag{9}$$

Triglycerides T

$$\frac{dT}{dt} = -\frac{V_T E_L T}{K_T + T} \, , \tag{10}$$

Total fatty acid F

$$\frac{dF}{dt} = -\kappa_2 \frac{dT}{dt} \, . \tag{11}$$

4 Numerical approach for estimating the parameters

The model described in Section 3 consists of a system of i coupled, nonlinear ordinary differential equations and was solved in MATLAB®. The function `ode15s` was used to advance in time the system

$$\frac{dy}{dt} = F(t, y),$$

where t is time, y is the solution vector containing each of the i variables we are solving for (namely L, C, etc) and F is the vector of nonlinear functions for the derivatives. The relative and absolute tolerances were set to 1×10^{-6}. Due to the stiff nature of the equations (the integration must be performed over a wide range of parameter values as a part of the optimisation routine), `ode15s` was used instead of the more usual `ode45` as it uses a variable order backward differentiation formula to perform the time stepping.

The model outputs, $\mathbf{y}$, were then used to fit a series of parameters to experimental data collated by Kim et al. [6], where the experimental data is denoted as $\mathbf{y}^{\mathrm{exp}}$. However, the experimental data was only available for k variables (where $\mathsf{k} \leqslant \mathsf{i}$) and at a limited set of time points, denoted t^*. The optimisation was therefore performed over these k variables and the model outputs were linearly interpolated to the same time points as the experimental data (these interpolated variables are denoted $\mathbf{y}^*$). A relative residual sum of squares over the variables given in the experimental data was used to compute the scalar error metric

$$E = \sum_{t^*} \sum_{k} \frac{(\mathbf{y}_k^* - \mathbf{y}_k^{\mathrm{exp}})^2}{\mathbf{y}_k^{\mathrm{exp}}} \, .$$

This error metric, E, was then minimised by the function `fminsearch` from the Global Optimisation toolbox. Parameters fitted in this fashion are explicitly listed in Table 3.

With the limited experimental data only some of the parameters could be fitted in this way. Other parameters have had their values fixed at appropriate values with reference to the literature [6].

5   Results

Kim et al. [6] choose to split their validation into two "phases", which correspond to disjoint sequential time periods. The experimental data is essentially divided into these two phases. Phase 1 is governed by biological processes during which quantities of bacterial cells continue to change and enzymes are produced, whereas phase 2 is primarily chemical in nature.

We follow the division into parts; however, we divide the modelling process into three stages, the results for each stage being described separately. The time periods for these stages are not disjoint: stage 1 fits data from just the phase 1 time period, whereas stages 2 and 3 fits data over the entire ripening period. The parameters from the simulations are given in Table 3.

Table 3: Parameter Values: (a) Fixed values [6, etc.]; (b) Stage 1 values found by fitting; (c) Stage 2 values found by fitting; (d) Lipolysis values found by fitting.

K_A	0.207	mg g^{-1}	(a)
K_B	1.15	mg g^{-1}	(a)
k_1	0.005	day^{-1}	(a)
k_2	0.0235	day^{-1}	(a)
Y_x	1.04×10^9	cfu mg^{-1}	(a)
ζ	1.08	dimensionless	(a)
κ_2	1.0	dimensionless	
K_ℓ	4.2322×10^4	mg g^{-1}	(b)
k_ℓ	0.2388	day^{-1}	(b)
μ_{LA}	3.692×10^{-7}	mg cfu^{-1} day^{-1}	(b)
μ_m	4.7295×10^3	day^{-1}	(b)
V_b	9.4449×10^{-12}	mg U^{-1} day^{-1}	(c)
V_f	$1.9752 \times 10^{-11} \times (e_{1,0})^{-1}$	mg U^{-1} day^{-1}	(c)
α_1	$1.2972 \times e_{1,0}$	U cfu^{-1}	(c)
μ	0.5151	U cfu^{-1}	(c)
k_3	0.00256	day^{-1}	(d)
K_T	1.5537	mg g^{-1}	(d)
V_T	2.864×10^{-9}	mg U^{-1} day^{-1}	(d)
μ_L	2.2119×10^{-4}	U cfu^{-1}	(d)

In stage 1 (relating to phase 1), the values of K_1, k_1, μ_{LA} and μ_m were used as fitting parameters as described in Section 4. In stage 2, these parameters were fixed at the values from stage 1, with V_b, V_f, α_1 and μ then used as fitting parameters. Unlike Kim et al. [6], who set specific initial conditions at the start of phase 1 and 2 simulations separately, we use a single set of initial conditions for both stages corresponding to the data from Kim et al. [6]. This is given in Table 4. A third stage fits the lipolysis reaction. As there is limited experimental data, some of the quantities and parameters have not been fitted. Lactic acid has not been modelled further here (although there

Table 4: Initial conditions used for both phases.

X	3.692×10^9	cfu g^{-1}
L	15.366	mg g^{-1}
α	1	mg g^{-1}
E_1	$3.692 \times 10^9 \times e_{1,0}$	U g^{-1}
E_2^o	0.023	U g^{-1}
A	258	mg g^{-1}
C	2.07059	mg g^{-1}
B	24.7013	mg g^{-1}
E_L	1	U g^{-1}
T	9.5	mg g^{-1}
F	0.5	mg g^{-1}

is additional consideration in Section 7).

5.1   Stage 1

The results from fitting the model to the phase 1 data collated by Kim et al. [6] are shown in Figure 2. Phase 1 is characterised by biological activity, as the cells consume lactose to reproduce and also die and undergo lysis. This part of the ripening process is described by equations (1)–(4). Figure 2a shows that at around day two, the cell population has increased from the initial value. This corresponds to a large decrease in the concentration of lactose, as shown in Figure 2b. Then, the cells, having almost no food left, die at a steady rate. As seen in Figure 2, this matches well to the experimental data.
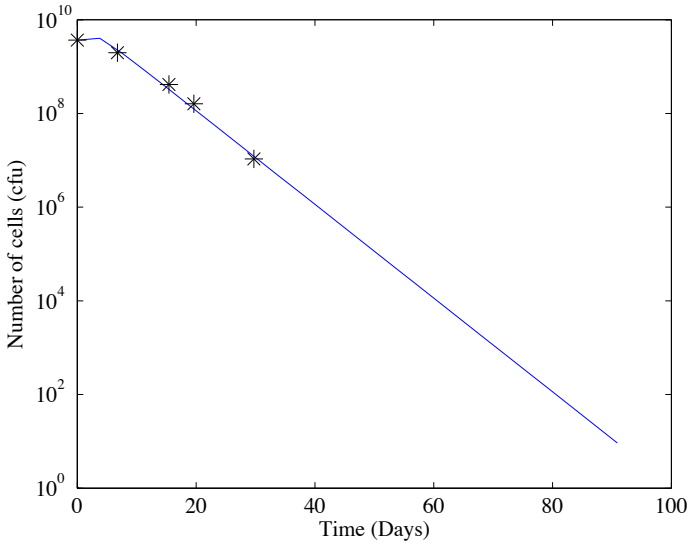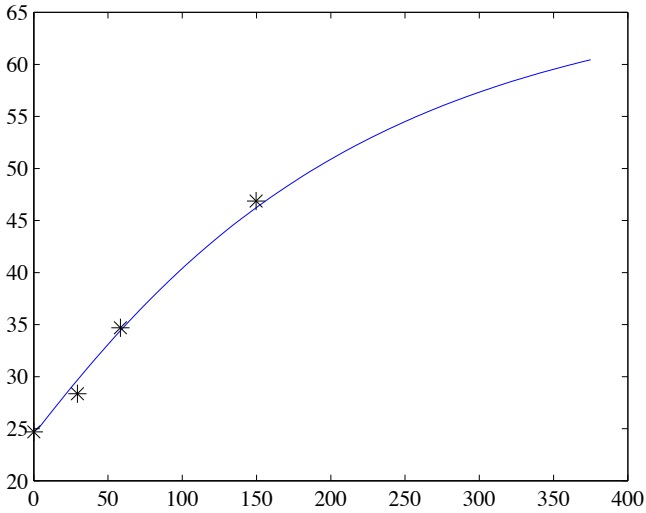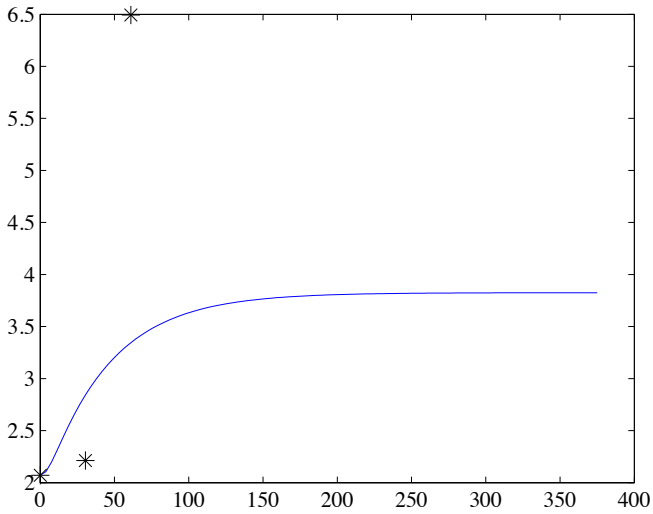
(a) Cells
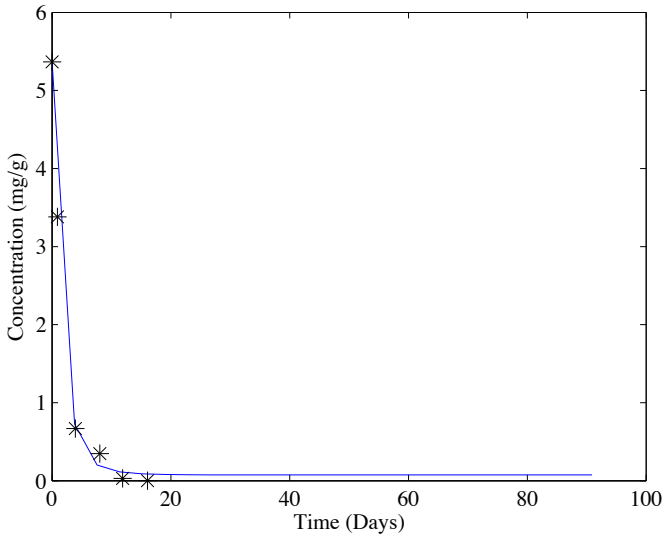
(b) Lactose

Figure 2: Stage 1 results. The symbols correspond to experimental data, and solid lines to the model output.

(a) Peptides

(b) Amino acids

Figure 3: Stage 2 results. The symbols correspond to experimental data, and solid lines to the model output.

5.2   Stage 2

Phase 2 occurs when the majority of the cell population has died, due to the lack of available lactose (as detailed in Section 5.1). The only remaining processes that can occur in the model are chemical; namely the breakdown of various proteins into peptides and then amino acids and the breakdown of fats (triglycerides) into fatty acids. These processes continue until the cheese is considered mature. We consider just the former proteolysis process here described by equations (4)–(8). The latter lipolysis process is stage 3, considered in Section 5.3. Figure 3 shows the results from fitting the model in the second stage. The proteinase E_1 is expressed in terms of multiples of $e_{1,0}$ the initial proteinase activity per cell at the beginning of the ripening process, which has units U/cfu. Figure 3a shows that the model outputs match the experimental data very well for the peptide concentration, namely, the peptide concentration increases over time, as the casein degrades. Unfortunately, the fit for the amino acids was poor (Figure 3b). There are only three data points and one appears to be a major outlier so the information available is rather limited.

5.3   Stage 3: Lipolysis

We performed a similar exercise to obtain the parameters required for the evolution of fatty acids. This occurs over the whole ripening period. Our model was fitted to fatty acid data presented by Marsili [9]. Given that equations (9)–(11) can be solved independently, we fixed the parameters fitted in Stages 1 and 2. Also, as we had no data on the proportionality constant κ_2 this was fixed at one. If another value were to be used, then there would be a corresponding change in V_T.

We see that our model system is able to achieve good agreement with the experimental data (Figure 4). The parameters used for the fitting, and their resulting values, are listed in Table 3.
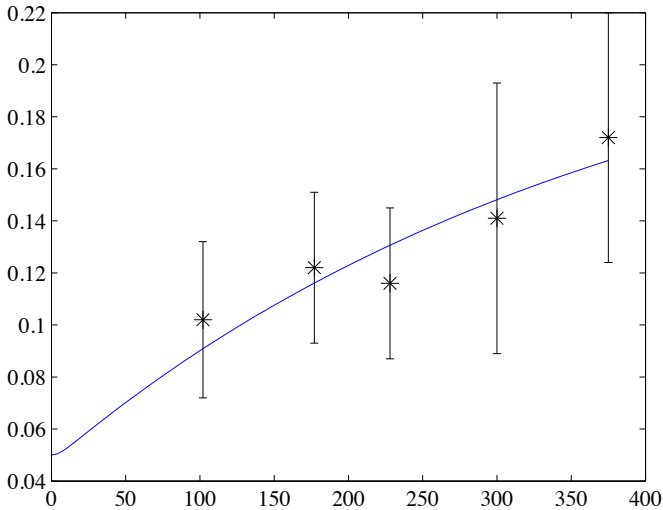
Figure 4: Stage 3: Lipolysis results. The symbols and error bars correspond to experimental data for fatty acids, and the solid line to the model output.

6   Statistical analysis

To aid the investigation into the ripening of cheddar cheese, Fonterra provided two data sets to be analysed. Four properties of cheese were recorded: the fat, water, protein and salt contents.

For the first set of data, samples were taken from cheese blocks, immediately after their production. Measurements are made by accurate laboratory tests (for which the results are returned after some delay) and by immediate NIR spectrometer readings (taken in duplicate with only the mean provided for analysis). Further measurements were made approximately one week later. These used the average of about eight instrument tests, with the focus being on the differences over a week. The cheese blocks are wrapped and stored for ripening (in the usual way) for the week between measurements.

Most initial samples were within the general specifications for cheddar which
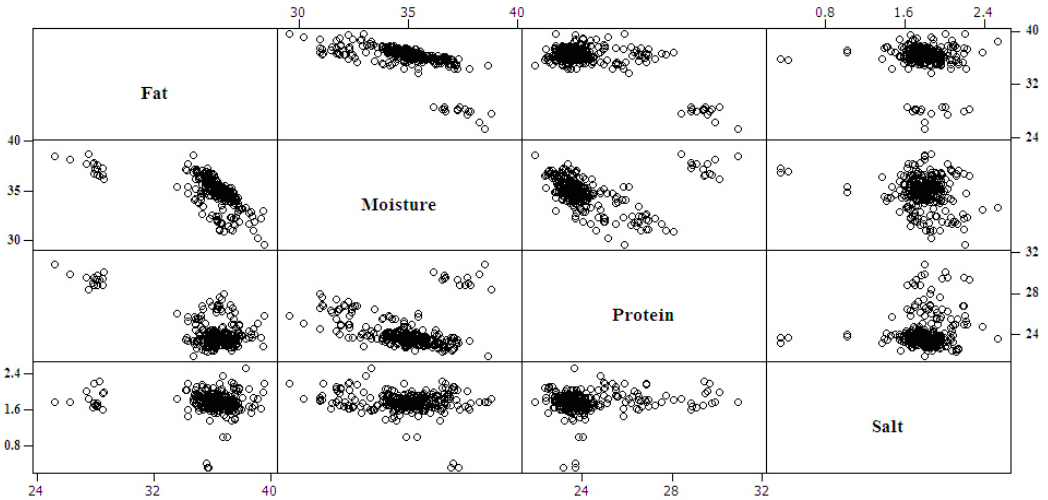
Figure 5: Fat, Moisture, Protein and Salt Content (%), in freshly produced cheese blocks.

are 33% to 39% fat, 0% to 38% moisture, and 1.4% to 2% salt. However, the samples were not homogeneous as Figure 5 shows. This figure indicates that different sites may have been working to different specifications. Furthermore, there is a group of observations for a low fat cheddar that display different characteristics to the remaining samples. After discussion with industry representatives, these observations (Fat < 30%) were removed from the analysis and do not appear in subsequent calculations and figures.

The differences over a week show negative correlation between the changes in fat and moisture, and also between the changes in moisture and protein. Thence we have a positive correlation between fat and protein changes. The Pearson correlation coefficients and their respective p-values are shown in Table 5 and are annotated on Figure 6. The Pearson coefficient measures the linear correlation between two data sets and takes values between 1 (total positive correlation) and −1 (total negative correlation) with 0 being no correlation. The p-values that lie between 0 and 1 give the statistical significance
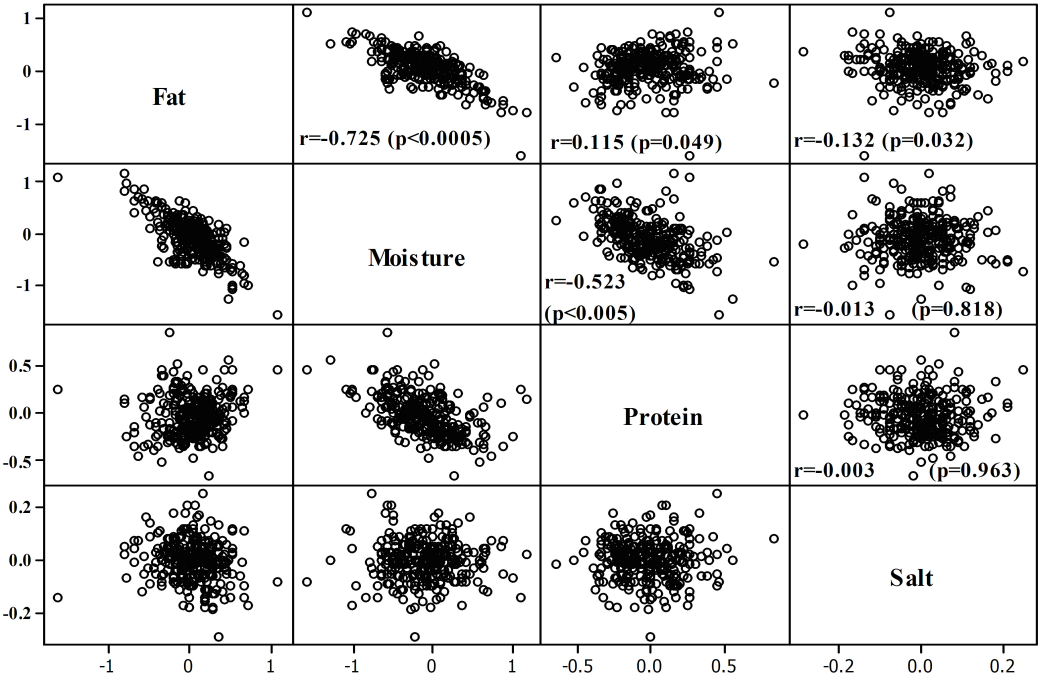
Figure 6: Changes in composition during the first week of ripening (freshly produced–one week later).

of the correlation coefficient with small values being more significant.

For the second set of data, samples were taken from the cheese blocks immediately after production and each sample was then stored in cool conditions (separate to the cheese blocks). Measurements were made on the samples over a period of about a month. There were at least two NIR spectrometer observations per day taken for samples at different sites and different spectrometers. The changes in composition over time are shown in Figure 7 for one instrument.

The proportions of moisture declined over time in all samples, while the proportions of fat and protein increased. The degree to which this occurred
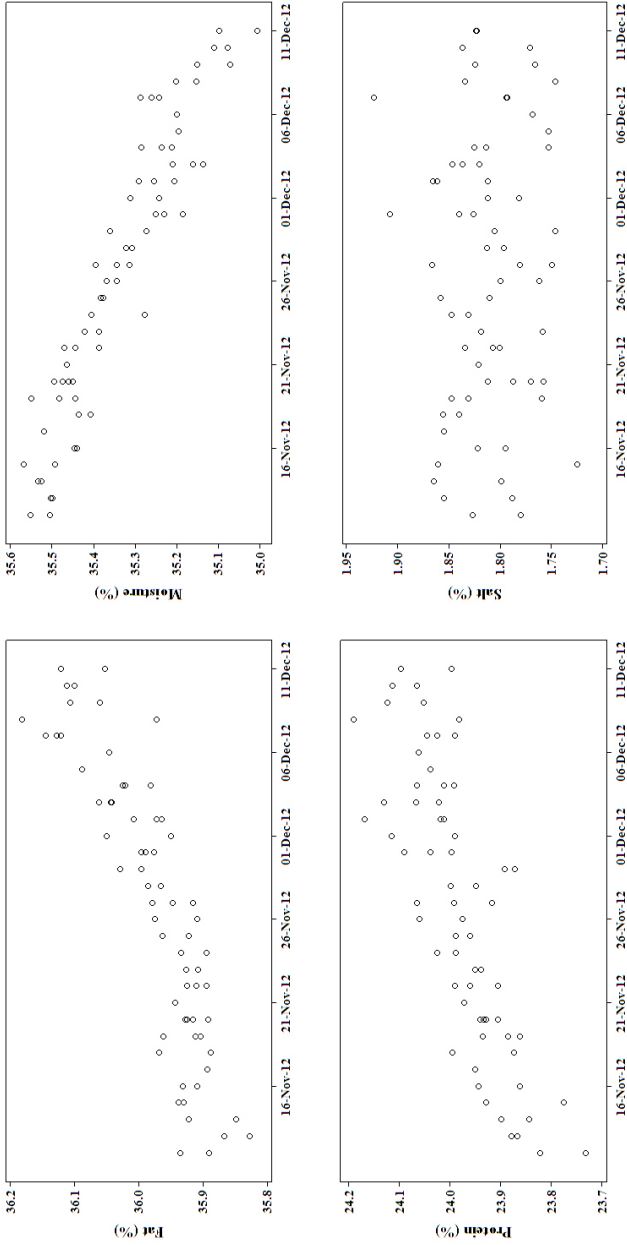
Figure 7: Changes in percentage of fat, moisture, protein, and salt over time, by instrument.

Table 5: The correlations of the differences in components for cheddar ($\geqslant 30\%$ fat) over a week. Cell contents: Pearson correlation (p-value)

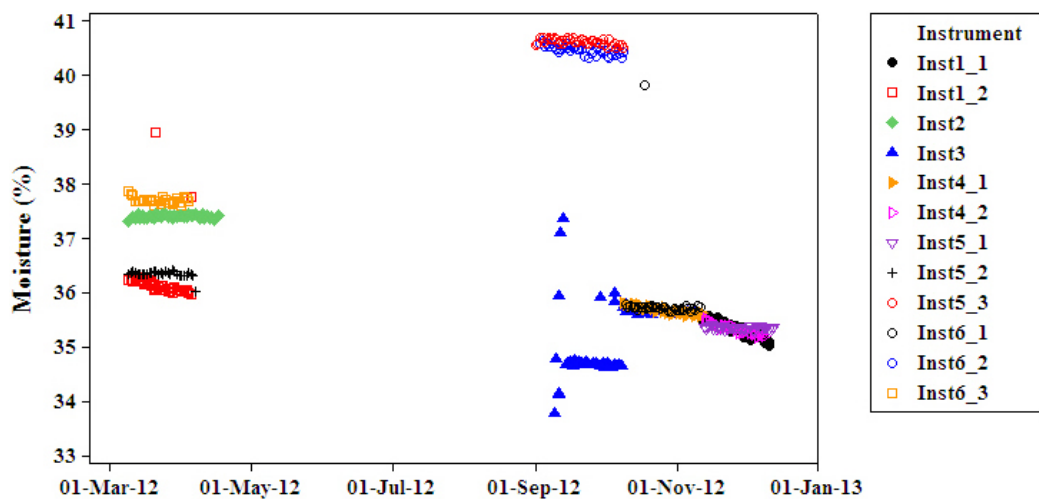|	Fat	Moisture	Protein
Moisture	$-0.725\,(< 0.0005)$		
Protein	$0.115\,(0.049)$	$-0.523\,(< 0.0005)$	
Salt	$-0.132\,(0.023)$	$-0.013\,(0.818)$	$-0.003\,(0.963)$

Figure 8: Changes in moisture (%) over time by instrument

was masked for some data sets by the observations not being for precisely the same grade of cheese. The significant differences between the instruments reflected the different specifications of cheddar being produced rather than differences in instrumentation. Overlaying all the instruments' measurements for moisture over time illustrates why the data could not be considered homogenous (Figure 8).

The ratio of protein to fat is an important factor in the making of cheddar cheese. The time series plot in Figure 9 shows this measurement for one of the instruments. The pH is also an important specification for cheese; however,
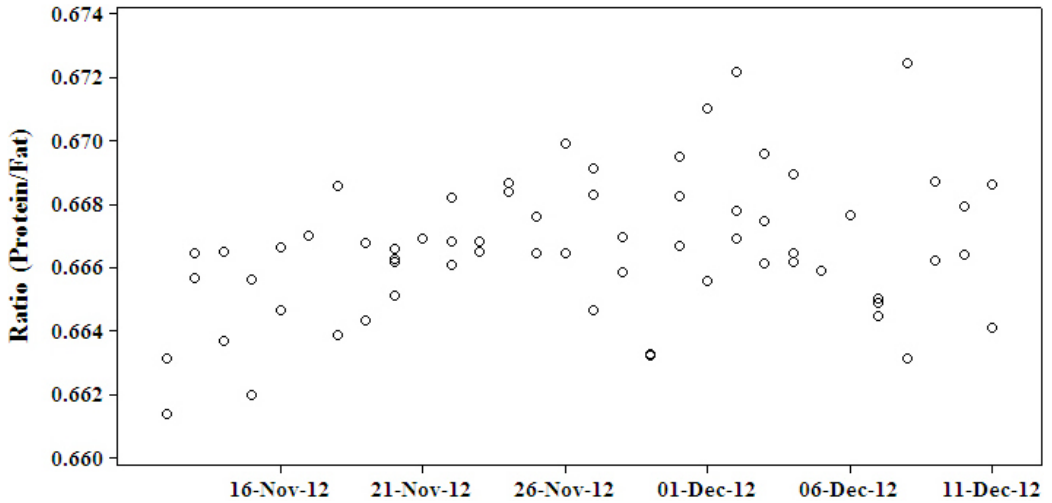
Figure 9: Changes in the Protein/Fat ratio for one instrument.

this measurement was not provided in either data set.

Due to the non-homogeneity of the data and insufficient information, these data were of limited assistance in modelling the properties of the ripening cheese. It would be helpful to have measurements taken over the whole period of ripening of the cheese. It is also desirable to have a more standardised testing with time series collected from the same site with the same cheese specification and, perhaps, the same cheese.

7 Mathematical modelling of pH changes in ripening cheese

The acidity of a cheese can have an important effect on properties such as texture and taste [8]. For many types of cheese it changes with time [3], and can be measured relatively easily [5]. However, cheddar cheese is often

reported as demonstrating no significant change in pH after production [11], despite the complex series of chemical reactions continuing as the cheese ripens.

During the MISG, the study of cheese acidity was conducted in parallel with other investigations and some aspects of the pH model differ from those in the eventual main model (Sections 3–5). Parameters had to be assigned values from limited information and ideally would be found more accurately. Some initial conditions used for the illustration are for raw milk. Quantities have been scaled. However, this preliminary (toy) model demonstrates how pH can be incorporated into the modelling process, and also indicates the kind of behaviour to expect. The model captures the general evolution of acidity in cheddar and is consistent with only modest changes in pH after production.

7.1   The model

The primary effector of pH change is the starter bacteria supplied [5]. These bacteria consume the lactose present in the milk (and later the curd) as a nutrient both to subsist and to proliferate. The consumption of lactose corresponds to a change in pH. Lactose is broken into glucose and galactose. Glucose then is further broken down into lactic acid:

$$C_6H_{12}O_6 \longrightarrow 2\,CH_3CHOHCOOH\,. \tag{12}$$

As in Section 3, we model the bacteria cells, X, the lactose, L, and the present concentrations of lactic acid, α. In this section we also include other weak organic acids, W, some of which are produced from the lactic acid. The dynamic evolution of these biological and chemical species over time is

modelled by

$$\frac{dX}{dt} = \mu_m^* \, f(pH) \, \frac{LX}{K_\ell + L} - k_\ell^* X \left(1 - \frac{L}{K_2 + L} \right) , \tag{13}$$

$$\frac{dL}{dt} = -\frac{\mu_m^* \, f(pH)}{Y_x} \, \frac{LX}{K_\ell + L} - \mu_{LA}^* \frac{LX}{K_2 + L} , \tag{14}$$

$$\frac{d\alpha}{dt} = 2 \frac{dL}{dt} - d_\alpha \alpha , \tag{15}$$

$$\frac{dW}{dt} = d_\alpha \alpha . \tag{16}$$

As mentioned previously there are some differences with the equations of Section 3. The earlier model assumed that pH did not change significantly, whereas here it can change with potential effects upon the ripening process. The starred parameters in equations (13) and (14) have a similar role to their unstarred counterparts in equations (1)–(3); however, they differ in detail. Here the cell reaction constant μ_m has been replaced by the product of a modified cell reaction constant μ_m^* and a function which depends on the acidity $f(pH)$. The reduction in growth rate with acidity is easily observed [4, 10], and is described by

$$f(pH) = 1 - 10^{(pH_{min} - pH)}, \tag{17}$$

following the model used by Presser et al. for the growth rate of *E. coli* [10]. It can be seen that when the pH value reaches pH_{min}, $f = 0$ and no further growth can occur. Cell death (at base rate k_ℓ^*) is driven by starvation when the concentration of lactose is sufficiently low ($L \sim K_2$). The model includes separate terms for consumption of lactose for reproduction and subsistence, following Kim et al. [6], albeit with a difference in functional form. If fermentation is assumed homolactic (cf. equation (12)), then each glucose molecule produces two molecules of lactic acid and so a factor of two has been included in equation (15). (A fuller analysis of the differences in molecular weights between reagents and products may lead to further scaling factors similar to those included in the equations of Section 3.) The slow break

down of lactic acid into weaker organic acids, through a range of complex processes, is modelled by a simple exponential decay with decay constant d_α. The effects of non-starter lactic acid bacteria and the breakdown of fats into fatty acids has not been explicitly modelled.

The pH is calculated here using standard equilibrium considerations. The two relevant equilibria (one each for lactic acid and for the weak acids) are

$$K_{aa} = \frac{[A^-][H^+]}{[HA]}, \quad K_{aw} = \frac{[W^-][H^+]}{[HW]}. \tag{18}$$

Here K_{aa} is the dissociation constant for lactic acid (which is known), and K_{aw} is the dissociation constant for the weaker organic acids. Defining the concentrations of dissociated acids $[A^-] = a$ and $[W^-] = b$ necessitates that $[H^+] = a + b$, $[HA] = \alpha - a$ and $[HW] = W - b$. Substituting these values into the equilibria and rearranging,

$$0 = K_{aa}(\alpha - a) - a(a + b),$$
$$0 = K_{aw}(W - b) - b(a + b). \tag{19}$$

This system of equations can be solved for a and b using the multivariable Newton's method, allowing the pH to be calculated according to

$$pH = -\log_{10}([H^+]) = -\log_{10}(a + b). \tag{20}$$

Buffering effects of the cheese have not been considered here and these may be significant [5].

Equations (13)–(16) have steady states $(X, L, \alpha, W) = (0, L, 0, W)$, corresponding to a lack of bacterial action and all lactic acid having decayed into weaker acids. However, as decreases in bacterial cell population are associated with starvation (low values of L), the steady states $(0, 0, 0, W)$ are more realistic, representing a full conversion of lactose into weaker acids. This corresponds to a fully aged cheese. In this simple model, the conditions for a constant bacterial cell population are obtained by equating the right-hand

side of equation (13) to zero. Apart from the trivial solution $X = 0$, we find an approximate solution for $pH = pH_{min}$ with $L \gg K_2$. This indicates the bacterial population is steady when the pH has reached a level to halt bacterial growth if sufficient lactose is present for subsistence. This continuation of bacterial effect does match general understanding of cheese making; however, a quantitative study suggests that it is more appropriate for some bacterial species to use a straight exponential decay instead of some kind of starvation effect [1].

Determination of a complete set of parameter values and initial conditions is difficult. Values corresponding to raw milk (with added bacterial culture) were taken as a starting point. The model (equations (13)–(16)) captures both the initial bloom of bacteria and the long-term behaviour. Milk has essentially no lactic acid and a known pH, allowing for equation (20) with $a = 0$ to be solved for b, and thus the necessary initial concentration of weak acid W_0 to be determined via equations (19). The initial population of bacteria is chosen to be a value much smaller than the maximal population which is scaled to be unity. Although the typical lactose concentration in milk is well-known, it is more difficult to quote a figure for cheese given its solid nature, especially in the context of relating lactose to the acid equilibria concentrations equations (18). Instead, a value is used which results in a reasonable amount of acid production in the long term. The dissociation constant for weak acids, k_W, is chosen to be one tenth of that for lactic acid, which is a known value. The required concentration for survival of bacteria is defined as half the required concentration for proliferation, with these values chosen so that they are significantly lower than the initial lactose concentration. The consumption of lactose for proliferation is defined to be 250 times the consumption of lactose for survival, a choice which might seem unrealistic but agrees (to the same order) with the ratio of these parameter values used by Kim et al. [6] in their own ODE model of cheese maturation. Table 6 lists parameter values, and Table 7 gives the initial conditions. Quantities are approximate and scaled within this toy model and so units are not included. Dimensions are similar to those in Section 3.

Table 6: Example set of parameter values predicting appropriate bacterial activity in the conversion from milk into cheese for the equations (13)–(16).

Parameter	Definition	Value
μ_m^*	proliferation rate of bacteria	5
k_ℓ^*	death rate of starving bacteria	0.02
c_1	consumption rate of lactose—proliferation	2.5×10^{-5}
μ_{LA}^*	consumption rate of lactose—survival	1×10^{-7}
K_1	Michaelis–Menten constant for proliferation-driven consumption	1×10^{-7}
K_2	Michaelis–Menten constant for survival-driven consumption	5×10^{-8}
d_α	decay of lactic acid into weaker acids	0.05
K_{aa}	dissociation constant for lactic acid	1.75×10^{-5}
K_{aw}	dissociation constant of weaker acids	1.75×10^{-6}
pH_{min}	Minimum pH for starter bacteria growth	5.5

Table 7: Initial conditions for the model defined by equations (13)–(16).

Variable	Definition	Value
X_0	initial concentration of bacteria	0.01
L_0	initial concentration of lactose	5×10^{-6}
α_0	initial concentration of lactic acid	0
W_0	initial concentration of weak organic acids	2.223×10^{-7}

7.2   Results for the pH model

Numerical simulation of the equations (13)–(16) with the parameters and initial conditions from Tables 6 and 7 demonstrates the expected behaviour for cheddar cheese. The bacterial cell population quickly rises to the maximum value that can be supported by the cheese before dying off as the lactose present declines. Although lactic acid is being constantly produced by the consumption of the lactose, the decay of lactic acid into weaker acids counteracts this effect. A sharp initial decrease in pH, which lasts a day or two in this model, could be taken to represent the initial stages of cheese making that occur prior to ripening (although the actual conditions are somewhat different before the production of cheese blocks). The initial phase in which the pH decreases sharply is shown in Figures 10(a) and 11(a). Thereafter the pH remains roughly constant over the whole ripening period, showing only a slight decrease while the bacterial effect is active, and then a slight increase as the lactic acid is fully converted into weaker acids. This trend is shown by the dark line in Figures 10(b) and 11(b), with those figures also illustrating the effect of varying selected parameters.

Varying the value of pH_{min} changes the level of acidity at which bacterial cell reproduction can take place and could arise by the use of different starter bacteria for cheese production. As illustrated by Figure 10(b), with the same initial amount of lactose, all the simulations eventually result in the same steady pH, corresponding to a full conversion of the lactose into lactic acid and then into weak acid. However, for bacteria that require a higher pH to thrive, this conversion occurs very slowly. (It is still not complete after a full year for $pH_{min} = 6$ in the figure.) These trends in pH could correspond to those observed for cheeses other than cheddar, with slow gradual increases or decreases during ripening after the initial rapid effect of the starter bacteria.

Rehman et al. [11] varied the initial amount of lactose used in cheddar cheese production and examined the time evolution of lactose and pH. Their results are based on the mean from two sets of cheesemaking trials. In each set of trials there were three kinds of cheese which were made with low (0.25%),
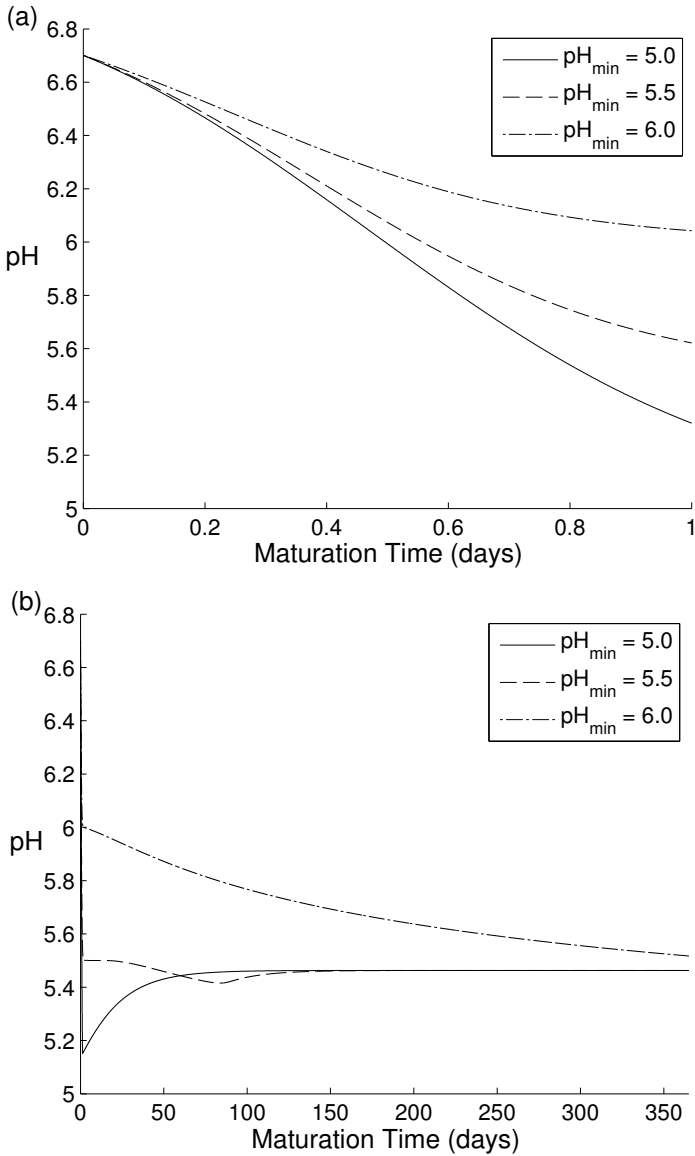
Figure 10: Changes in acidity as the cheese matures, for several different values of pH_{min} which might represent different types of bacteria: (a) changes over one day; (b) changes over one year.
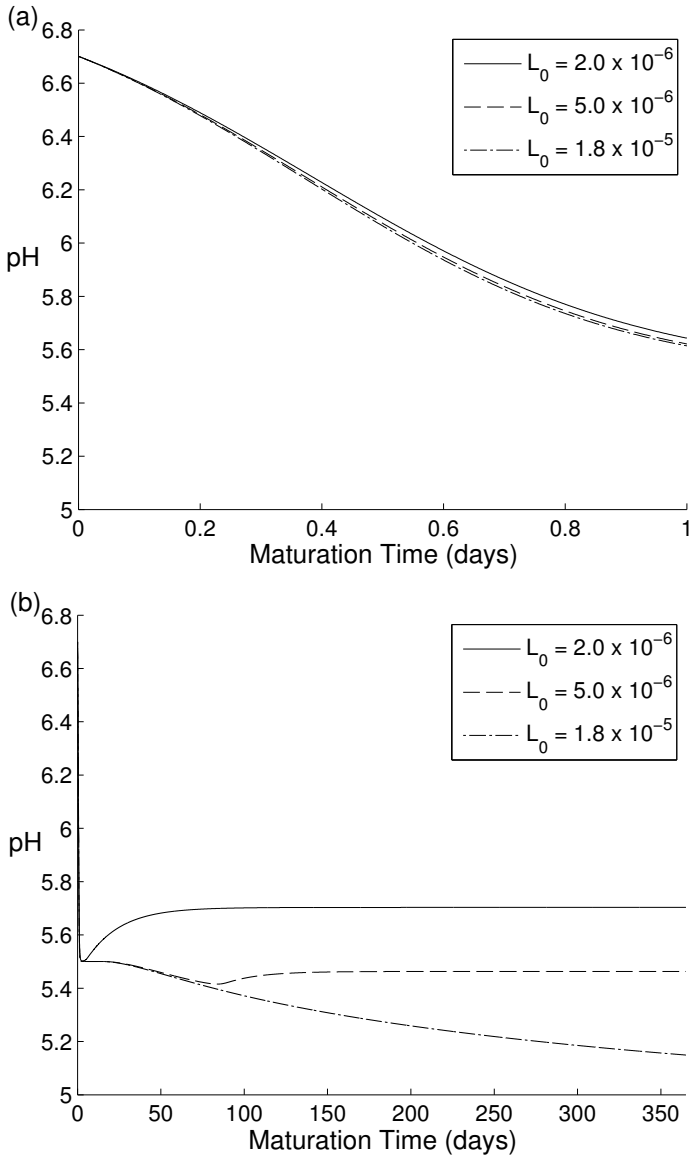
Figure 11: Changes in acidity as the cheese matures, for several different values of L_0 corresponding to different lactose levels in the milk used for production: (a) changes over one day; (b) changes over one year.

standard (0.61%) and high (2.20%) initial concentrations of lactose. Rehman et al.'s measurements were taken at 30 day intervals and the pH results which we now describe are shown on Figure 2 in their paper. The initial pH values were similar, with low, standard, and high initial lactose concentrations corresponding to 5.32, 5.28 and 5.25, respectively. However, during the 180 day experiment these values diverged. The pH of the low-lactose cheese trends slightly upward, although, the data is quite variable. The initial pH 5.32 increases to nearly pH 5.4 at 90 days but then sharply declines between 90 and 120 days to close to the initial value and is slightly below that value for the 150 day measurement. The final result at 180 days is just above pH 5.4.

In Rehman et al.'s standard and high-lactose cheese the results are monotonic. In standard cheese the pH is essentially constant, the marginal decline in the data going from pH 5.28 to pH 5.25 in 180 days. For the high-lactose cheese, the pH drops steadily to just below 4.8.

Figure 11 illustrates how a similar variation in initial lactose content affects our pH model. It is appropriate to compare Rehman et al.'s experimental results with Figure 11(b) as the experimental measurements were taken after the block of cheese has been formed. As in Rehman et al.'s experiments, the equilibrium pH now varies according to the initial amount of lactose (L_0), with more lactose resulting in a larger total acid production and thus lower final pH. The pH model results shown in Figure 11(b) are broadly consistent with those of Rehman et al., in that there is a relatively constant pH for standard cheese, a steady decrease in pH when initial lactose is increased and an increase when initial lactose is decreased. The simulated pH in all cases begins from approximately the same value, matching experimental observations.

The present pH model has considered both the short-term and long-term effects of the starter lactic acid bacteria, with starkly different timeframes between the initial bacterial bloom and the cheese's maturation. The evolution of the model species over a shorter timeframe reveals that the initial bacterial bloom up to some maximal population occurs over a timeframe of approximately a day, matching general knowledge for cheddar manufacture [12]. The corresponding

acidification is shown in Figure 11(a).

8   Discussion and conclusions

We considered a number of the aspects of cheese ripening modelled during the MISG. Further to the description here, some members of the MISG team considered how sensory test data could be incorporated with the measurements from the NIR instruments. Their idea involved investigating the use of machine-learning algorithms to map the data obtained from the NIR measurements to the sensory testing. It was envisaged that this may allow Fonterra to predict the grade of cheese that will be produced at an early stage. Currently, these ideas have not been implemented; however, this is an avenue for future investigation.

The model of Section 3 captured some of the key features in the ripening of cheese. It shows promise in comparison with experimental data in Section 5. In particular, our model shows excellent agreement with experimental data observed in Figures 2, 3(a) and 4. The precision of the amino acids fitting in Figure 3(b) was limited by the data available. To verify the model predictions for amino acids, further experimental data need to be collected, preferably over a longer period with more than three data points and with an estimate of experimental error.

While some data were made available by Fonterra, these data were of limited use in the modelling exercise due to their non-homogeneity and the insufficient information provided. Some aspects of the cheese data provided were investigated in Section 6. Further data are required for a deeper statistical analysis as well as being required to obtain better fits to the model parameters. Accurate measurements for each of the main variables in our model over the timescale of cheese ripening would provide greater verification of the mathematical model developed in this work.

The pH model explored in Section 7 captures another aspect of the cheese-

making process. As changes in acidity in ripening cheddar are small the effect may be modest. However, in other cheeses there are greater changes in pH during the ripening process and such a pH model may have a greater importance.

References

[1] D. Dawson and J. Feagan, Bacteriology of cheddar cheese: A study of starter organisms in manufacture and maturing. *J. Dairy Res.* **24**, 210–224, 1957. doi:10.1017/S002202990000875X M29

[2] M.A. Drake, R.E. Miracle and D.J. McMahon, Impact of fat reduction on flavor and flavor chemistry of Cheddar cheeses. *Journal of Dairy Science* **93**, 5069–5081, 2010. doi:10.3168/jds.2010-3346 M12

[3] M. Fenelon and T. Guinee, Primary proteolysis and textural changes during ripening in cheddar cheeses manufactured to different fat contents. *Int. Dairy J.* **10**, 151–158, 2000. doi:10.1016/S0958-6946(00)00040-6 M25

[4] R. Hutkins and N. Nannen, pH homeostasis in lactic acid bacteria. *J. Dairy Sci.* **76**, 2354–2365, 1992. doi:10.3168/jds.S0022-0302(93)77573-6 M27

[5] M. Johnson, *Cheese pH—what's behind the rise and fall?* Technical Report 14, Wisconsin Center for Dairy Research, December 2002. https://www.cdr.wisc.edu/sites/default/files/pipelines/2002/pipeline_2002_vol14_04.pdf M25, M26, M28

[6] J.K. Kim, K. Starzak, G.W. Preckshot, R. Marshall and R.K. Bajpai, Critical reactions in ripening of cheeses—a kinetic analysis. *Appl. Biochem. Biotech.* **45**, 51–68, 1994. doi:10.1007/BF02941787 M5, M6, M11, M14, M15, M16, M27, M29

[7] B.A. Law and A.S. Wigmore, Microbial proteinases as agents for accelerated cheese ripening. *Int. J. Dairy Tech.* **35**, 75–76, 2007. doi:10.1017/S0022029900032751 M10

[8] R. Lawrence, L. Creamer, and J. Gilles, Texture development during cheese ripening. *J. Dairy Sci.* **70**, 1748–1760, 1987. doi:10.3168/jds.S0022-0302(87)80207-2 M25

[9] R. Marsili, Monitoring chemical changes in cheddar cheese during aging by high performance liquid chromatography and gas chromatography techniques. *Journal of Dairy Science* **68**, 3155–3161, 1985. doi:10.3168/jds.S0022-0302(85)81221-2 M7, M19

[10] K. Presser, D. Ratkowsky, and T. Ross, Modelling the growth rate of Escherichia coli as a function of pH and lactic acid concentration. *Appl. Environ. Microbiol.* **63**, 2355–2360, 1997. http://aem.asm.org/content/63/6/2355.full.pdf M27

[11] S. Rehman, D. Waldron, and P. Fox, Effect of modifying lactose concentration in cheese curd on proteolysis and in quality of Cheddar cheese. *Int. Dairy J.* **14**, 591–597, 2004. doi:10.1016/j.idairyj.2003.11.008 M26, M31

[12] T. Singh, M. Drake, and K. Cadwallader, Flavour of cheddar cheese: A chemical and sensory perspective. *Compr. Rev. Food Sci. F.* **2**, 166–189, 2003. doi:10.1111/j.1541-4337.2003.tb00021.x M34

[13] S. Toelstede and T. Hofmann, Kokumi-active glutamyl peptides in cheeses and their biogeneration by Penicillium roquefortii. *Agric. Food Chem.* **13**, 3738–3748, 2009. doi:10.1021/jf900280j M11

Author addresses

1. **Winston L. Sweatman**, Centre for Mathematics in Industry, Institute of Natural and Mathematical Sciences, Massey University, Auckland, NEW ZEALAND.
   mailto:w.sweatman@massey.ac.nz

2. **Steven Psaltis**, School of Mathematical Sciences, Queensland University of Technology, AUSTRALIA.
   mailto:steven.psaltis@qut.edu.au

3. **Steven Dargaville**, School of Mathematical Sciences, Queensland University of Technology, AUSTRALIA.
   mailto:dargaville.steven@gmail.com

4. **Alistair Fitt**, Oxford Brookes University, UNITED KINGDOM.
   mailto:afitt@brookes.ac.uk

5. **Tony Gibb**, Adelaide Advanced Engineering, AUSTRALIA.
   mailto:agibb@351halifax.com

6. **Brodie Lawson**, School of Mathematical Sciences, Queensland University of Technology AUSTRALIA.
   mailto:b.lawson@qut.edu.au

7. **Kaye Marion**, RMIT, AUSTRALIA.
   mailto:kaye.marion@rmit.edu.au