Comparison of Several Forms of Equations for Predicting Cheddar Cheese Yield from Milk Composition

Craig A. Moore

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Food Science Commons

Recommended Citation
Moore, Craig A., "Comparison of Several Forms of Equations for Predicting Cheddar Cheese Yield from Milk Composition" (1984).
All Graduate Theses and Dissertations. 5312.
https://digitalcommons.usu.edu/etd/5312
COMPARISON OF SEVERAL FORMS OF EQUATIONS FOR PREDICTING CHEDDAR CHEESE YIELD FROM MILK COMPOSITION

by

Craig A. Moore

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

Nutrition and Food Sciences

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1984
ACKNOWLEDGEMENTS

I wish to thank Dr. R.J. Brown, my major professor, for his guidance and encouragement throughout my course work and research. His advice and help were extremely valuable to me.

I am indebted to Dr. C.A. Ernstrom for his insightful questions and constructive criticism during this work.

I'm grateful to Dr. D.V. Sisson for his encouragement and willingness to help me during my study.

I would also like to thank Joe Heap at the Glacier Mountain Cheese Company for the work he did that made this research possible.

Much gratitude is owed to my dear wife, Jody, for her help and continued support of my efforts.

I also want my fellow students in the laboratory for their interest and suggestions.

Craig A. Moore
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>REVIEW OF LITERATURE</td>
<td>2</td>
</tr>
<tr>
<td>Factors Affecting Cheddar Cheese Yields</td>
<td>2</td>
</tr>
<tr>
<td>Formulas for Predicting Cheese Yields</td>
<td>9</td>
</tr>
<tr>
<td>Milk Pricing Systems for Cheesemaking</td>
<td>12</td>
</tr>
<tr>
<td>METHODS AND PROCEDURES</td>
<td>17</td>
</tr>
<tr>
<td>Milk and Cheese Source</td>
<td>17</td>
</tr>
<tr>
<td>Milk Analysis</td>
<td>17</td>
</tr>
<tr>
<td>Milk and Cheese Weights</td>
<td>17</td>
</tr>
<tr>
<td>Cheese Moisture Analysis</td>
<td>17</td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>18</td>
</tr>
<tr>
<td>RESULTS</td>
<td>20</td>
</tr>
<tr>
<td>DISCUSSION AND CONCLUSIONS</td>
<td>27</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>31</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table                                    Page
1. Residual sums of squares for the eight models . . . . . . . . 22
2. Mean deviations for the eight models . . . . . . . . . . . . . . 23
3. Correlation coefficients (r) between actual
   and predicted yields of the eight models . . . . . . . . . . . . 24
LIST OF FIGURES

Figure

1. Regression of predicted cheese yields vs. actual cheese yields for Model 1

Page

26

Craig A. Moore, Master of Science
Utah State University, 1984

Major Professor: Rodney Jay Brown
Department: Nutrition and Food Sciences

This study was conducted to evaluate several forms of equations for predicting Cheddar cheese yields based on the fat and protein content of milk and moisture content of cheese. Production and quality control data from a Cheddar cheese plant for one entire year was used. This included the pounds of milk that went into each vat of cheese, yield of cheese from each vat, cheese moisture from each vat, and fat and protein percentages of the milk.

Seven models were derived to predict the yield of Cheddar cheese. The seven models were statistically fitted to the data by applying the Marquardt non-linear least squares method of iteration. These were compared with the commonly used Van Slyke and Price formula, with casein estimated as a percentage of total protein. The differences among the eight models were small.

(40 pages)
ABSTRACT

Comparison of Several Forms of Equations for Predicting Cheddar Cheese Yield from Milk Composition

by

Craig A. Moore, Master of Science
Utah State University, 1984

Major Professor: Rodney Jay Brown
Department: Nutrition and Food Sciences

This study was conducted to evaluate several forms of equations for predicting Cheddar cheese yields based on the fat and protein content of milk and moisture content of cheese. Production and quality control data from a Cheddar cheese plant for one entire year was used. This included the pounds of milk that went into each vat of cheese, yield of cheese from each vat, cheese moisture from each vat, and fat and protein percentages of the milk.

Seven models were derived to predict the yield of Cheddar cheese. The seven models were statistically fitted to the data by applying the Marquardt non-linear least squares method of iteration. These were compared with the commonly used Van Slyke and Price formula, with casein estimated as a percentage of total protein. The differences among the eight models were small.

(40 pages)
INTRODUCTION

Cheese yield is defined as the weight of cheese manufactured from a given weight of milk. Many formulas have been developed to predict cheese yield from milk composition. Most of these are based on the fat and casein percentages in milk (18). A formula derived by Van Dam and Janse (18) is based on fat plus solids-not-fat in the milk. Another formula proposed by Schulz and Kay (18) is based on fat and total nitrogen in the milk.

The cheese yield prediction formula receiving the widest acceptance was derived by Van Slyke and Price (59). It predicts cheese yield from the fat and casein percentages in milk and the moisture content of the cheese.

A modified Van Slyke and Price formula is being used by cheese plants to set prices for the milk they buy based on its cheese yielding capacity (10).

The objective of this study was to derive several formulas to predict the yield of Cheddar cheese from milk fat and protein percentages and compare them with the commonly used Van Slyke and Price formula. These formulas could then help individual Cheddar cheese plants develop more equitable pricing systems for milk.
Factors Affecting Cheddar Cheese Yields

Three fundamental factors control Cheddar cheese yield. They are the composition of milk, the percentage of milk constituents lost during manufacture, and the amount of moisture retained in cheese (6,33).

Casein and fat content of milk

The milk components most important to cheese yield are fat and casein (6). A direct linear relationship between the amount of fat and casein in milk and Cheddar cheese yield has been demonstrated (45). This relationship forms the basis of many formulas that predict the yield of Cheddar cheese.

Calcium in milk

The amount of calcium in milk is another factor affecting Cheddar cheese yields (53). This is because milk salts, especially calcium, are essential for the formation of a rennet curd (6). Holt and Muir (29) observed that soluble calcium concentrations in milk vary significantly during the year. They reach a minimum about the time that the casein to fat ratio reaches a peak, thus clouding the importance of both effects.

Milk handling

Many variables influence the percentage of milk constituents lost in Cheddar cheese manufacture. Volume losses during handling reduce cheese yield (40). Physical abuse of fat globules during milk handling lowers yields. Abuse can occur by pumping with inadequately sized pumps, improper operation of separators, and allowing milk to splash into vats.
Such practices disrupt the fat globule membrane causing release and loss of fat (40). Van Slyke and Price (59) indicated that fat loss can in part be recovered by slow warming and stirring of milk to work the fat back into an emulsion.

Microbial quality of milk

The microbial quality of raw milk used in cheesemaking has an impact on cheese yield. Psychrotrophic bacteria actively produce extracellular proteinases. When high numbers of such bacteria are present in raw milk they may produce sufficient levels of enzymes to cause significant breakdown of casein (16). Allauddin, et al. (2) noted that when large numbers of psychrotrophs were present in milk, cheese yield was reduced. Hicks, et al. (26) indicated that yield loss resulted from both lipid and protein degradation. This loss was caused by psychrotrophic Bacillus and Pseudomonus present in milk.

Milk storage

Phelan (44) showed that milk composition changes as it is stored. Refrigerated storage significantly decreased curd-forming properties of milk because of the growth of psychrotrophic bacteria (15). If native casein has been sufficiently altered, milk clotting enzymes will not work effectively (53). Microbial proteinases can inhibit milk clotting by masking reactive sites on the casein molecule or by inactivating milk clotting enzymes. Weak sets and incomplete gel formation allow more solids to be lost in the whey. Cousins, et al. (16) noticed that while pasteurization kills Gram negative psychrotrophs, it does not inactivate their enzymes. Therefore, heat resistant proteinases continue to solubilize casein, thus reducing the yield of cheese.
Milk pasteurization

Davis (17) reported a slight increase in yield for cheese made from pasteurized milk over that made with raw milk. Phelan (44) indicated that heat precipitation facilitates the incorporation of whey proteins into cheese curd. Increased yield is also due to the enveloping effect of precipitated protein on the fat globules (18). Brown and Ernstrom (12) found 75°C to be more satisfactory for a 30 minute heat treatment than 60°C or 90°C for trapping whey protein in the curd. Teese (56) observed that pasteurization at higher temperatures and shorter times favored a higher yield.

Culture media

Salji and Kroger (51) suggested that direct to the vat cultures produce greater cheese yields than phosphate bulk cultures. Hicks, et al. (27) observed greater yields using direct to the vat sets and enriched ammoniated whey base cultures compared to citrate base or skimmilk media.

Culture proteinase activity

Richardson (48) stated that starter cultures with higher ratios of proteinase negative cells to proteinase positive cells give greater cheese yields. He reasoned that less casein is digested in both the culture tank and cheese vat when proteinase activity is low, thus reducing casein losses in whey.

Milk clotting enzymes

The effect of milk-clotting enzymes on cheese yield is unclear. Emmons, et al. (20) observed no significant yield difference in cheese

trials of pepsin combined with chymosin EC.3.4.4.3 compared to chymosin alone. Using chymosin, 50/50 chymosin-pepsin, and pepsin, Fox and Walley (23) produced Cheddar cheese with no major difference in fat and protein losses in whey. However, Emmons, et al. (20) reported slightly higher fat and protein levels in whey and slightly lower yield of cheese made with bovine pepsin compared to cheese made with chymosin. Using rennet from Mucor pusillus Lindt, Nelson (38) claimed that, in comparison with chymosin, Cheddar cheese yields did not differ significantly. Robertson and Gilles (49) obtained similar results using the same clotting enzyme. Antila and Aapola (3) found that with enzyme from M. pusillus Lindt, more fat was in the whey and the yield of cheese was slightly lower than with chymosin. Cheese made with Endothelia parasitica enzymes produced a lower yield than with the control (19).

Puhan and Irvine (47) compared Cheddar cheese yields produced with Bacillus subtilis to those made with chymosin. He concluded that high proteolysis from bacterial proteases lowered the yield by 10%. Some investigators measured 22% more non-protein nitrogen in whey using M. pusillus extracts than calf chymosin (53). Using M. miehei they measured 17.8% more non-protein nitrogen in whey compared to chymosin.

The proteinases produced from E. parasitica had higher proteolytic activity than enzymes from the Mucor species (pusillus and miehei) while the animal types (chymosin, bovine pepsin, and porcine pepsin) were significantly lower (53). Calf chymosin was the least proteolytic and porcine pepsin was the most proteolytic of the animal types.

Addition of coagulant

Sellars (53) recommends maximum dispersibility of the coagulant
both by adequate dilution with slightly chlorinated (5 ppm) water at pH 6.5 and uniform dispersement to allow maximum efficiency of the enzymes. Catalytic activity is lost when enzyme molecules are not evenly dispersed in the vat. Loss of catalytic activity usually results in weaker sets and therefore lower yields. Pearce (43) advises adding rennet directly, in a slow stream, to stirred milk. He believes this is the best way to prevent chlorine inactivation of rennet that can occur when rennet is diluted with chlorinated water.

Setting of vat

Failure to keep fat well distributed before setting results in accumulation of fat at the surface of the milk, most of which ends up in whey (59). Movement in the vat while milk is being set disrupts the milk gel, resulting in lower cheese yields (41). Fisk (22) reports that setting milk at a higher temperature reduces fat loss in whey, thus increasing cheese yields. Sellars (53) indicates that milk pH at setting affects enzyme clotting activity. Low enzyme activity weakens the milk gel resulting in lower yields. Porcine pepsin is the most sensitive milk clotting proteinase to pH's above 6.6 followed by bovine pepsin and calf chymosin. Microbial proteinases are least affected by high milk pH's.

Cutting of curd

Fisk (22) lost more fat in whey by cutting the curd fine than by cutting it coarse. He also lost 0.13% more fat in whey when the curd was cut soft than when it was cut hard. Bynum and Olson (13) obtained increased fat retention and cheese yield with high curd rigidity at cutting. Olson (40) indicated that cutting the curd too soft causes its
disruption and additional fat loss in whey. He also said the additional force and stresses created by cutting an overly firm coagulum may result in curd breakage and therefore reduce yields.

Heating the curd

Van Slyke and Price (59) stated that more fat is lost in whey by heating curd too rapidly or to too high a temperature.

Stirring the curd

Olson (40) recommended not stirring the curd for a few minutes after cutting. This permitted "healing" of curd surfaces by forming a membranous film, thus minimizing curd breakage. Fisk (22) discovered that stirring curd as the last of the whey is removed caused a larger fat loss in whey and a reduced cheese yield.

Cheddaring the curd

Excessive piling of curd, previous to cheddaring, increases fat loss (59).

Salting the curd

If curd is salted at a temperature above 32 C., fat is likely to extrude with the whey and be lost (59).

Milling the curd

Milling cheddared blocks exposes new curd surfaces and consequently fat globules. Dull knife cutting edges of curd mills cause increased abrasion and squeezing of cheddared curd (40). This results in erosion of fat globules and therefore lower cheese yields. Milling at too high a temperature or allowing the curd to become matted after milling
increases loss of fat (59).

Pressing the curd

Fisk (22) discovered that pressing the curd fast reduced yield because more fat was squeezed out of the curd.

Moisture retention

The amount of moisture retained in cheese directly affects cheese yields. The water content of cheese is independent of the water content of milk. The only significant factor in milk composition relating to cheese moisture is the casein to fat ratio (31). A high casein to fat ratio favors a higher moisture cheese (46,19).

Cheesemaking procedures have a strong impact on the amount of water retained in cheese. Fisk (22) found that as more rennet extract is added to cheese milk the percentage of moisture in cheese is increased. Setting milk at a higher temperature increases moisture retention (57). The increased moisture from using more rennet extract and a higher setting temperature occur because a tighter curd is formed. As curd is cut into smaller pieces, less moisture remains in the cheese (33,59). Cutting the curd hard causes a higher moisture cheese (7,60). If curd is heated rapidly after cutting, moisture retention increases (57). Rapid heating is accompanied by formation of a tough film on the outside of curd particles which seals moisture in (59). A high cooking temperature favors less moisture in cheese (33,60). Allowing the curd to remain in whey for long periods permits reabsorption of the whey and thus increases cheese moisture (57,60). If the whey is removed while the acid level is low, cheese moisture increases (22,60). By stirring the curd as the last of the whey is removed, less moisture is retained
Piling curd quickly after removing the whey increases cheese moisture (57). More frequent turning of curd blocks decreases moisture retention (33). By piling curd blocks higher, more moisture remains in cheese (22,59). As the cheddaring process is prolonged, cheese moisture increases (60). If the curd is soaked in water previous to salting, more water is retained (57,60). The addition of greater amounts of salt to the curd increases moisture loss (22,59). Pressing the curd quickly favors water retention (22). High temperature and low humidity in the curing room increase moisture loss from cheese as it ages (57,60).

Formulas for Predicting Cheese Yields

Many formulas have been derived by many different workers to predict Cheddar cheese yield based on various milk components.

A. Formulas based on fat content:

1. Van Slyke and Price (58) simplified their basic formula by substituting the following for casein: percent casein = 0.4(F - 3) + 2.1 to obtain the following formulas:
   a. Yield = 2.7fat  
   b. Yield = 1.1fat + 5.9  
   c. Yield = 2.3fat + 1.4

B. Formulas based on fat and casein:

1. Babcock, et al. (5):
   Yield = 1.1fat + 2.5casein  
2. Van Slyke and Price (59):
   Yield = 1.63(casein + fat)  
3. McDowell (36) (using Van Slyke's data):
   Yield = 1.4(casein + fat) + 1.04
4. McDowell (36) (using his data):
   a. Yield = 1.22(casein + fat) + 2.32
   b. Yield = 1.07fat + 2.35casein

5. Shelton (54):
   Yield = (F - 4F/100) + (C - 4C/100 + 22C/100)2.26
   This was derived on the following premises:
   a. Loss of 4% of the fat in whey.
   b. Loss of 4% of the casein in whey.
   c. Retention in the cheese of non-casein solids-not-fat equivalent to 22% of the casein.
   d. Cheese moisture content equivalent to 126% of the solids-not-fat retained.

6. The above formula was simplified by McDowell (18):
   Yield = 0.96fat + 2.67casein

C. Formula based on total solids or fat plus solids-not-fat:
   Van Dam and Janse (18) emphasized the simplicity of the following formula:
   Yield = fat + 1/3 solids-not-fat.

D. Formulas based on fat and total nitrogen (or total protein):
   1. Schulz and Kay (18):
      Yield = net fat + 1.8P

      where:
      P = protein content of the milk, assuming that 75% of this goes into the cheese.

   2. Phelan (44) listed the following formulas from New Zealand:
      a. Yield = 2.18P + 1.17F
      b. Yield = (1.62P + 1.21F)(wt. milk lbs.)/100
3. Banks, et al. (6):

\[ \text{Yield} = 1.58F + 1.23P \]

E. Formulas based on milk fat and protein in milk and water and salt in cheese:

Bergman and Joost (18) favored the following formulas:

1. \( \text{Yield} = 0.91F + 0.77P + 0.48 + W(0.77P + 0.48)/(100 - W) \)

2. \( \text{Yield} = 91F + 77P + 40/(100 - S + W) \)

where:

\( F \) = percentage of fat in milk.
\( P \) = percentage of protein in milk.
\( W \) = percentage of water in cheese.
\( S \) = percentage of salt in water.

F. Formulas based on fat and casein in milk and moisture content of cheese:

1. Joost, et al. (18):

\[ \text{Yield} = 0.866F + 0.752P + 0.460 + W(0.75P + 0.460)/(100 - W) \]

where:

\( F \) = percent fat in milk.
\( P \) = percent protein in milk.
\( W \) = percent moisture in the fat free fraction of the cheese.

2. Van Slyke and Price (59):

\[ Y = (0.93F + C - 0.1)1.09/(1 - W) \]

where:

\( Y \) = Kg. Cheddar cheese per 100 Kg. milk.
\( F \) = percent fat in the milk.
\( C \) = percent casein in the milk.
\( W \) = Kg. water per Kg. cheese.
3. Price (45) modified the previous formula, giving it more flexibility:

\[
Yield = \frac{(F \times R) + (C - 0.1)}{1.09/TS}
\]

where:

- \( F = \% \) fat in milk.
- \( R = 100 - \% \) fat lost during cheesemaking/100
- \( C = \% \) casein in milk.
- \( TS = 100 - \% \) moisture of cheese/100

Milk Pricing Systems for Cheesemaking

Originally milk was bought and sold solely on the basis of volume (11). This system encouraged skimming and watering down of milk (21). In the 1890's milk began to be paid for on the basis of fat. This was made possible by the development of the Babcock test in the United States and the Gerber test in Europe, which tested for milkfat.

During the first part of this century the price of milk was completely based on its fat content (10). Butter, churned from milkfat, could be placed in frozen storage for extended periods. Butter was considered the "balance wheel" of the dairy industry because, during the flush season, milk could be separated and the cream made into butter for sale during the slack seasons. With skim milk having little value, butter stabilized milk prices throughout the year.

Cheese makers realized that cheese yields were dependent on both the casein and fat content of milk (11). The Jacobsen formula, available at that time, described the relationship between fat and casein in milk. It said that percent casein in milk varies in the same direction as percent fat, but only 40% as much (24). Therefore, with
milk prices based solely on fat content, cheese makers found it more profitable to buy low-fat milk because the casein cost per pound was lower in that milk (11). This spurred the establishment of minimum standards of fat in cheese with 50% fat in the dry matter being required for Cheddar (50).

World War II brought about a shortage of butter which was coupled with a substantial improvement in the quality of margarine (10). Many people became accustomed to margarine and milk fat lost its ability to absorb all the value of milk.

As the value of solids-not-fat increased relative to the value of fat there was increasing pressure to adjust the price of milk for the variation in solids-not-fat (24). It was realized that paying for milk on a straight butterfat basis was an inadequate pricing plan since it assumed that solids-not-fat increase at the same rate as fat content (55,1). Pricing programs were established based on a standard value per hundredweight of milk testing 3.5% fat with a fat differential that was added or subtracted for each 0.1% above or below 3.5% (24). The Federal Milk Marketing Administration and most plants not under Federal Milk Marketing Orders now use this pricing system, with the exception of California where milk pricing is based on its fat and solids-not-fat content (21).

The fat differential pricing system permits the dairy farmer to receive more money for his milk if he adds water to it (21). Although it is illegal for milk producers to directly add water to the milk they sell, they can add water by breeding and management practices (11). This pricing system has encouraged dairymen to do just that.

Herrington, et al. (25) observed that total non-fat-milk solids had
declined almost half a percent between 1892 and 1960 in New York dairy herds and that casein content had dropped from 2.56% to 2.35% in the skim milk fraction. Comparison of the results of a 1971 nationwide survey by Wilcox, et al. (61) on the composition of milk of the major dairy breeds with the 1945 data of Overman (42) demonstrated there were marked decreases in solids-not-fat percentages in the milk from these breeds during that 25 year period. Hoover, et al. (30) indicated in 1971 that in the preceding 15 years the national average fat content of milk declined from 3.86% to 3.68%. Schultz (52) stated that average milkfat content had changed from 4% in 1950 to 3.65% in 1972. Johnson (32) reported the following average decreases in milkfat in the United States: 1945, 3.98%; 1955, 3.84%; 1965, 3.70%; and 1970, 3.66%. He added that the decrease has been associated with a shift to the low-fat breeds.

Fat differential pricing has served the fluid milk industry well because demand over the past several years has been for lower fat milk products and the consuming public has not desired to pay more for higher solids-not-fat in their milk (21).

However, this system has not served the manufacturing industry well and has caused particular problems for the cheese industry (21). This is because it has encouraged lower solids in milk which has lowered cheese yields. In some cases, cheese plants have paid more for milk under this pricing system than the value of cheese they've been able to make from it (10).

This problem has encouraged many suggestions for alternate pricing schemes. Hillers, et al. (28) advocated milk pricing methods which deemphasize the fat component and place greater emphasis on the protein
or solids-not-fat components. Ladd and Dunn (34) proposed a two component pricing formula based on fat and protein or solids-not-fat, arguing that producers with low protein to fat ratios in their milk receive a disproportionately large share of milk payments. Brog (8) favored a fat-protein pricing model as did McGann and O'Connell (37) because of the simple and accurate testing methods that are now available for these two components. Johnson (32) suggested a pricing plan with a base fluid price per 100 lbs. of milk and a differential for both fat and protein, saying that protein is milk's most valuable component. The base fluid price would be different for three different milk classes. Brog (9) introduced the following formula for pricing milk for cheesemaking:

\[
\text{Price} = 2.6010AB + 1.0625CD + 1.1265AD + 0.1806E + 4.1905F
\]

where:

A = pounds of protein in 100 pounds of milk.
B = price per pound of 40 pound Cheddar blocks.
C = pounds of fat in 100 pounds of milk.
D = price per pound of butter.
E = price per pound of whey cream butter.
F = price per pound of whey powder.

Johnson (32) said that a problem with using a protein differential in milk pricing is that protein and solids-not-fat are not readily separable and that their value differs among manufactured uses. He suggested another way to approximate the value of milk components is from wholesale prices of dairy products. Brog (9) stated if one can predict the amount of cheese that can be obtained from a given quantity of milk, with protein and fat as estimators, it should be relatively
simple to equate cheese prices to producer milk values.

Ernstrom (21) proposed a pricing program based on the cheese yielding capacity of milk. It pays producers a fixed price per pound of cheese that their milk is predicted to produce. This prediction is based on the fat and protein content of the milk and the moisture content of cheese, using the Van Slyke formula (59). This pricing program is being successfully used by many cheese plants (11).

Milk Analysis

The percentage of fat in each milk sample was determined by the Babcock method (20). Milk samples were tested for percentage protein using the official A.O.A.C. protein-dye binding method (4).

Milk and Cheese Weights

Milk weights were measured with a calibrated dip stick in the cheese vat. Cheese weights were taken by weighing the curd before hooping. Then the blocks of cheese were weighed after pressing. This was done so that any curd not put into hoops could be added into the final cheese weight by calculating the percentage of weight lost by hooping and pressing the curd.

Cheese Moisture Analysis

Moisture in the cheese was determined by the official A.O.A.C. rapid screening method in a 130 C oven (4).
METHODS AND PROCEDURES

Milk and Cheese Source

Cheese was made at a commercial cheese plant during the period from January 1, 1981 to December 31, 1981. Weights of milk per cheese vat and cheese yield per vat were recorded. Analysis of milk fat and protein and cheese moisture were done in the cheese factory's laboratory.

Milk Analysis

The percentage of fat in each milk sample was determined by the Babcock method (39). Milk samples were tested for percentage protein using the official A.O.A.C. protein-dye binding method (4).

Milk and Cheese Weights

Milk weights were measured with a calibrated dip stick in the cheese vat. Cheese weights were taken by weighing the curd before hooping. Then the blocks of cheese were weighed after pressing. This was done so that any curd not put into hoops could be added into the final cheese weight by calculating the percentage of weight lost by hooping and pressing the curd.

Cheese Moisture Analysis

Moisture in the cheese was determined by the official A.O.A.C. rapid screening method in a 130°C oven (4).
Statistical Analysis

The collected data included a total of 560 cheese vats. The data was used to evaluate each of eight models. The models are:

**Model 1** (a modified Van Slyke and Price formula with casein estimated at 78% of total protein)

\[ Y_1 = \frac{(0.93F + 0.78P - 0.1)1.09}{(1 - W)} \]

**Model 2** (same as Model 1 except parameter replaces fixed casein percentage of protein and coefficients are all determined by iteration)

\[ Y_2 = \frac{(aF + bP - c)d}{(1 - W)} \]

**Model 3** (same as Model 2 except parameter \( d \) is absent)

\[ Y_3 = \frac{(aF + bP + c)}{(1 - W)} \]

**Model 4** (same as Model 3 except parameter \( c \) is absent)

\[ Y_4 = \frac{(aF + bP)}{(1 - W)} \]

**Model 5** (same as Model 4 plus casein:fat ratio controls amount of fat recovered in cheese)

\[ Y_5 = \frac{(0.78P/F)}{abF + cP}/(1 - W) \]

**Model 6** (same as Model 5 plus parameter \( d \))

\[ Y_6 = \frac{(0.78P/F)}{abF + cP + d}/(1 - W) \]

**Model 7** (same as Model 6 except parameter \( a \) is removed)

\[ Y_7 = \frac{(0.78P/F)}{(aF + bP + c)}/(1 - W) \]

**Model 8** (same as Model 7 except parameter \( c \) is absent)

\[ Y_8 = \frac{(0.78P/F)}{(aF + bP)}/(1 - W) \]
In all eight models:

\[ Y = \text{Kg cheese per 100 Kg milk.} \]

\[ F = \text{Percent fat in milk.} \]

\[ P = \text{Percent protein in milk.} \]

\[ W = \text{Kg moisture per Kg cheese.} \]

The parameters derived by Van Slyke and Price were used in Model 1. Marquardt non-linear least squares was used to fit Models 2 through 8 to the data. The Marquardt method performs a least squares fit of data to a function. The function can depend on any reasonable number of parameters and be non-linear. Data can have an arbitrary number of independent variables. Optimal function parameter values were found, starting from a set of initial guesses, to minimize the residual sum of squares (50). Marquardt's method is a compromise between Gauss-Newton and steepest descent and is most useful when the parameter estimates are highly correlated (35).

The parameter values in Van Slyke's formula were used as starting points for the iterations of the first model. Starting points for iterations of the other seven models were based on these values.
RESULTS

The results in Model 1 were derived by Van Slyke and Price:

Model 1

\[ Y_1 = \frac{(aF + 0.78P - b)c}{1 - W} \]

\[ Y_1 = \frac{(0.93F + 0.78P - 0.1)1.09}{1 - W} \]

The results in Models 2 through 8 were obtained by applying Marquardt iteration to the data:

Model 2

\[ Y_2 = \frac{(aF + bP - c)d}{1 - W} \]

\[ Y_2 = \frac{(0.93F + 0.79P - 0.094)1.1}{1 - W} \]

Model 3

\[ Y_3 = \frac{(aF + bP + c)}{1 - W} \]

\[ Y_3 = \frac{(0.77F + 0.91P + 0.65)}{1 - W} \]

The parameters obtained in this and all subsequent models have no physical meaning.

Model 4

\[ Y_4 = \frac{(aF + bP)}{1 - W} \]

\[ Y_4 = \frac{(0.79F + 1.09P)}{1 - W} \]

Model 5

\[(0.78P/F)\]

\[ Y_5 = \frac{(ab \quad F + cP)}{1 - W} \]

\[ Y_5 = \frac{((2.23)(3.23) \quad F - 3.67P)}{1 - W} \]

Model 6

\[(0.78P/F)\]

\[ Y_6 = \frac{(ab \quad F + cP + d)}{1 - W} \]

\[ Y_6 = \frac{((2.16)(3.22) \quad F - 3.67P + 0.57)}{1 - W} \]
Model 7

\[ \begin{align*}
Y_7 &= (aF + bP + c)/(1 - W) \\
    &= (0.29F + 1.30P + 0.65)/(1 - W)
\end{align*} \]

Model 8

\[ \begin{align*}
Y_8 &= (aF + bP)/(1 - W) \\
    &= (0.30F + 1.50P)/(1 - W)
\end{align*} \]

The correlation coefficients between actual yields and predicted yields for the eight models were found by linear regression. Table 1 shows the residual sums of squares for the eight models.
Table 1. Residual sums of squares for the eight models

<table>
<thead>
<tr>
<th>Model</th>
<th>RSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $Y_1 = (0.93F + 0.78P - 0.1)1.09/(1 - W)$</td>
<td>99.3</td>
</tr>
<tr>
<td>2. $Y_2 = (0.93F + 0.79P - 0.094)1.1/(1 - W)$</td>
<td>83.2</td>
</tr>
<tr>
<td>3. $Y_3 = (0.77F + 0.91P + 0.65)/(1 - W)$</td>
<td>80.4</td>
</tr>
<tr>
<td>4. $Y_4 = (0.79F + 1.09P)/(1 - W)$</td>
<td>80.9</td>
</tr>
<tr>
<td>5. $Y_5 = ((2.23)(3.23)F - 3.67P)/(1 - W)$</td>
<td>79.2</td>
</tr>
<tr>
<td>6. $Y_6 = ((2.16)(3.22)F - 3.67P + 0.57)/(1 - W)$</td>
<td>78.8</td>
</tr>
<tr>
<td>7. $Y_7 = (0.29F + 1.30P + 0.65)/(1 - W)$</td>
<td>80.1</td>
</tr>
<tr>
<td>8. $Y_8 = (0.30F + 1.50P)/(1 - W)$</td>
<td>80.7</td>
</tr>
</tbody>
</table>
Table 2 shows the mean deviation of the predicted yield from the actual yield for the eight models.

Table 2. Mean deviations of predicted from actual yields for the eight models.

<table>
<thead>
<tr>
<th>Model</th>
<th>MD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( Y1 = (0.93F + 0.78P - 0.1)1.09/(1 - W) )</td>
<td>0.270</td>
</tr>
<tr>
<td>2. ( Y2 = (0.93F + 0.79P - 0.094)1.1/(1 - W) )</td>
<td>0.274</td>
</tr>
<tr>
<td>3. ( Y3 = (0.77F + 0.91P + 0.65)/(1 - W) )</td>
<td>0.257</td>
</tr>
<tr>
<td>4. ( Y4 = (0.79F + 1.09P)/(1 - W) )</td>
<td>0.263</td>
</tr>
<tr>
<td>( (0.78P/F) )</td>
<td></td>
</tr>
<tr>
<td>5. ( Y5 = ((2.23)(3.23) F - 3.67P)/(1 - W) )</td>
<td>0.262</td>
</tr>
<tr>
<td>( (0.78P/F) )</td>
<td></td>
</tr>
<tr>
<td>6. ( Y6 = ((2.16)(3.22) F - 3.67P + 0.57)/(1 - W) )</td>
<td>0.256</td>
</tr>
<tr>
<td>( (0.78P/F) )</td>
<td></td>
</tr>
<tr>
<td>7. ( Y7 = (0.29 F + 1.30P + 0.65)/(1 - W) )</td>
<td>0.257</td>
</tr>
<tr>
<td>( (0.78P/F) )</td>
<td></td>
</tr>
<tr>
<td>8. ( Y8 = (0.30 F + 1.50P)/(1 - W) )</td>
<td>0.262</td>
</tr>
</tbody>
</table>
Table 3 shows the correlation coefficients (r) between actual yields and predicted yields for the eight models. Figure 1 shows the correlation coefficient between actual yields and predicted yields for model 1.

The data used in developing these equations includes 560 data points. In regressing actual yields against predicted yields (Figure 1) for model 1, small dots are used to represent the data points. Many points lie nearly on top of each other, thus leaving the impression that there are fewer data points than actually are present.

Table 3. Correlation coefficients (r) between actual and predicted yields of the eight models.

<table>
<thead>
<tr>
<th>Model</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Y1 = (0.93F + 0.78P - 0.1)1.09/(1 - W)</td>
<td>0.655</td>
</tr>
<tr>
<td>2. Y2 = (0.93F + 0.79P - 0.094)1.1/(1 - W)</td>
<td>0.655</td>
</tr>
<tr>
<td>3. Y3 = (0.77F + 0.91P + 0.65)/(1 - W)</td>
<td>0.638</td>
</tr>
<tr>
<td>4. Y4 = (0.79F + 1.09P)/(1 - W)</td>
<td>0.647</td>
</tr>
<tr>
<td>5. Y5 = ((2.23)(3.23)F - 3.67P)/(1 - W)</td>
<td>0.653</td>
</tr>
<tr>
<td>6. Y6 = ((2.16)(3.22)F - 3.67P + 0.57)/(1 - W)</td>
<td>0.645</td>
</tr>
<tr>
<td>7. Y7 = (0.29F + 1.30P + 0.65)/(1 - W)</td>
<td>0.639</td>
</tr>
<tr>
<td>8. Y8 = (0.30F + 1.50P)/(1 - W)</td>
<td>0.647</td>
</tr>
</tbody>
</table>
Figure 1. Regression of predicted cheese yields vs. actual cheese yields for Model 1.
DIDCUSSION AND CONCLUSIONS

The purpose of this study was to derive several formulas to predict the percentage of Cheddar cheese from milk composition and to compare them with the Van Slyke formula in current use in one cheese factory (59). Seven new casein formulas were derived to predict Cheddar cheese yield.

The first model was the Van Slyke and Price formula with casein estimated as % of total protein. The correlation coefficient of actual yield versus predicted yields was 0.655. This value was equal to or slightly higher than those values obtained for each of the seven derived models.

The second model used 79 casein estimates of total protein with a parameter added to increase the increased flexibility. The computer found the best value of the parameter for this equation. The parameters of this model were found to be very close to many cheese producers using lower fat reconversions. Both models using 79 percent of the casein in their cheese. These parameters show that 79 percent of the total milk protein is casein. The values of parameters c and d, found by iteration, were also nearly equal to those values in Van Slyke and Price's formula. The r value for this model was very close to Model 1's r value.

The third model was the same as Model 2 except that coefficients in the step before of the equation were multiplied together. Although this resulted in a lower model, the r value was lower than Model 2's value. The parameters of these models through b had no physical meaning.
DISCUSSION AND CONCLUSIONS

The purpose of this study was to derive several formulas to predict the yield of Cheddar cheese from milk composition and to compare them with the Van Slyke formula in current use in one cheese factory (59). Seven new model formulas were derived to predict Cheddar cheese yield.

The first model was the Van Slyke and Price formula with casein estimated as 78% of total protein. The correlation coefficient of actual yields versus predicted yields was 0.655. This value was equal to or slightly higher than the r values obtained for each of the seven derived models.

The second model replaced the 0.78 casein estimate of total protein with a parameter, thus giving the model increased flexibility. The computer found the best fit of parameters for this equation. The parameters indicated that, in contrast with many cheese plants reporting lower fat recoveries, this plant recovered 93 percent of the milkfat in their cheese. These parameters also imply that 79 percent of total milk protein was casein. The values for parameters c and d, found by iteration, were also nearly equal to the values in Van Slyke and Price's formula. The r value for this model was very close to Model 1's r value.

The third model was the same as Model 2 except that coefficients in the top half of the equation were multiplied together. Although this resulted in a simpler model, the r value was lower than Model 2's value. The parameter values for Models 3 through 8 had no physical meaning. Efforts to put bounds on the parameters to give them physical meaning resulted in a less accurate fit to the data.
The fourth model modified Model 3 by eliminating parameter c. This not only simplified the equation but improved the r value of the model.

The fifth model was like the previous model except that the casein to fat ratio was taken into account. However, its r value was still slightly lower than Model 2's r value.

The sixth model added another parameter to Model 5. However, this merely complicated the equation and resulted in a lower correlation coefficient.

Model 7 changed the sixth model by multiplying parameters a and b together. Although this simplified the equation, the r value did not improve.

The eighth model modified Model 7 by eliminating parameter c. This model was the simplest of the four models that included the casein to fat ratio. Its r value was higher than Models 6 and 7, but lower than Model 5.

Models 5 through 8 differed from the first four models because they included the casein to fat ratio. However, this did not improve the r values of the equations. This likely occurred because the casein percentages of the milk samples were not known and were estimated at 78% of total protein. By estimating the percent casein at 78% of total protein, much of the variation in casein to fat ratios between milk samples was lost, thereby negating the value of including it in the equations.

Table 1 shows the residual sums of square for the eight models. Model 6 has the lowest residual sum of squares with 78.8 while Model 1 has the highest at 99.3. Table 2 shows the mean deviation of the predicted yield from the actual yield for each model. Model 6 has the
lowest mean deviation of 0.256 while Model 2 has the largest mean deviation at 0.274. Table 3 shows the correlation coefficients for the eight models. Models 1 and 2 have the highest r value with 0.655 while Model 3 has the lowest with 0.638.

Although there are differences between the residual sums of squares, the mean deviations, and the correlation coefficients of the eight models, none of the differences are large. Therefore, any of the eight equations will do a good job of predicting cheese yields relative to each other. Model 1 is recommended because of the physical meaning held by its parameters and its demonstrated accuracy relative to other cheese yield prediction formulas.

One of the other models would be very useful in a factory where fat recovery in the cheese is less than 93% or where the percent casein of total protein varies significantly from 78 percent. In that case, Model 4 is recommended due to its simplicity. The techniques outlined in this study are adaptable to any cheese plant. Thus, a cheese plant could derive parameters that would best predict cheese yield in that factory by using its own data. To do this, accurate records of the protein and fat content of the milk used in each cheese vat must be kept. Also, milk and cheese weights and cheese moistures must be kept for each vat. The data must be obtained for each vat rather than averaging the data so as to include all the variability between cheese vats.

Increasing numbers of cheese plants are setting the price they pay for milk on the expected cheese yielding capacity of the milk. Many plants have found good success by using the Van Slyke formula to predict cheese yields. The formulas derived in this study may make it possible for cheese makers to more accurately predict the cheese yielding
capacity of the milk they buy if their fat or protein recovery is different than that predicted by Van Slyke's formula. The Van Slyke and Price and the formulas derived in this study enable the Cheddar cheese industry to improve their cheese yield pricing systems and permit the dairy farmer to receive a more fair price for the milk he sells.

These formulas necessarily include the variation in cheese making procedures from vat to vat. They can be used to compare the expected yield of each vat with its actual yield. Then cheese making practices can be modified in low yielding vats to bring them more into line with expected yields.

Only the regression line (Figure 1) for model 1 was included in the results because the regression lines for all eight models were nearly identical.

The regression line of equation one demonstrates that the models tend to underestimate the yield of high solids milk. This probably occurs because a higher percentage of high solids milk comes from Jersey cows. The protein in Jersey milk averages 80% casein compared to 78% in Holstein milk (14). Holsteins are the predominant breed. Therefore, further work needs to be done to develop better equations for predicting Cheddar cheese yields from milk with very high solids.

Further study is recommended to develop a simple, accurate method for measuring casein in milk. This would eliminate the need to estimate casein from total milk protein. The direct measurement of casein would eliminate this important source of variation in Cheddar cheese yield prediction formulas. More accurate equations for predicting Cheddar cheese yield could then be developed by using casein directly rather than a constant percentage of total protein.
REFERENCES


46. Price, W.V. and L. Germain. 1931. Standardization of milk for the


