Modeling buffer capacity and pH in acid and acidified foods

Robert E. Price, Madyson Longtin, Summer Conley-Payton, Jason A. Osborne, Suzanne D. Johanningsmeier, Donald Bitzer, and Fred Breidt

Abstract: Standard ionic equilibria equations may be used for calculating pH of weak acid and base solutions. These calculations are difficult or impossible to solve analytically for foods that include many unknown buffering components, making pH prediction in these systems impractical. We combined buffer capacity (BC) models with a pH prediction algorithm to allow pH prediction in complex food matrices from BC data. Numerical models were developed using Matlab software to estimate the pH and buffering components for mixtures of weak acid and base solutions. The pH model was validated with laboratory solutions of acetic or citric acids with ammonia, in combinations with varying salts using Latin hypercube designs. Linear regressions of observed versus predicted pH values based on the concentration and p*K* values of the solution components resulted in estimated slopes between 0.96 and 1.01 with and without added salts. BC models were generated from titration curves for 0.6 M acetic acid or 12.4 mM citric acid resulting in acid concentration and p*K* estimates. Predicted pH values from these estimates were within 0.11 pH units of the measured pH. Acetic acid concentration measurements based on the model were within 6% accuracy compared to high-performance liquid chromatography measurements for concentrations less than 400 mM, although they were underestimated above that. The models may have application for use in determining the BC of food ingredients with unknown buffering components. Predicting pH changes for food ingredients using these models may be useful for regulatory purposes with acid or acidified foods and for product development.

Keywords: acid, base, acid foods, acidified foods, buffer capacity, buffer model, pH

Practical Application: Buffer capacity models may benefit regulatory agencies and manufacturers of acid and acidified foods to determine pH stability (below pH 4.6) and how low-acid food ingredients may affect the safety of these foods. Predicting pH for solutions with known or unknown buffering components was based on titration data and models that use only monoprotic weak acids and bases. These models may be useful for product development and food safety by estimating pH and buffering capacity.

1. INTRODUCTION

The market for acid and acidified foods and beverages in the United States includes fermented and acidified vegetables, sports drinks, carbonated beverages, fruit juices, salad dressings and sauces, salsas, and others. The U.S. acidified food regulations define acid and acidified foods as having an equilibrium pH value at or below pH 4.6 (FDA, 1979). A pH below 4.6 has been shown to be sufficient to prevent botulism in these food products (Ito & Chen, 1978). For many of these foods, weak organic acids are either naturally present, produced by fermentation, or added to maintain the pH. Most common food acids have p*K* values near or below pH 4.6 (CRC, 1995) and function as buffers in the pH range of most acid or acidified food products, typically pH 3.2 to 4. Buffering in acid or acidified foods is important in maintaining the equilibrium pH below 4.6 for the entire shelf life of these foods.

In addition to lowering pH and acting as buffers, the protonated forms of weak acids are antimicrobial (Breidt, 2006; Hosein, Breidt, & Smith, 2011; Russel, 1992) and have been shown to contribute to sour taste perception (Neta, Johanningsmeier, Drake, & McFeeters, 2009). For weak acid preservatives, uncharged protonated weak acids are relatively hydrophobic and can diffuse through bacterial membranes. Once inside of a bacterial cell, they dissociate in the cytoplasm, which typically has a pH around neutral (pH 6 to 7). This results in two problems for a bacterial cell in the presence of weak acids in acid or acidified foods where the external pH is 2 to 3 pH units below the cytoplasmic pH. First is the acidification of the cytoplasm of the cells, and second is the intracellular accumulation of weak acid anions (Russel, 1992). Therefore, the antimicrobial efficacy of weak acid preservatives, including sodium benzoate and potassium sorbate, as well as the acids present in most acidic foods and beverages, is dependent on the ratio of protonated to dissociated (acid anion)

JFDS-2019-1570 Submitted 9/25/2019, Accepted 1/31/2020. *Authors Price, Longtin, Conley-Payton, Johanningsmeier, and Breidt are with U.S. Department of Agriculture, Agricultural Research Service, SEA, Food Science Research Unit, NC State University, 322 Schaub Hall, Box 7624, Raleigh, NC 27695, U.S.A. Author Longtin is also with Department of Food, Bioprocessing and Nutrition Sciences, NC State University, 400 Dan Allen Drive, Raleigh, NC 27695, U.S.A. Author Osborne is with Department of Statistics, NC State University, 2311 Stinson Drive, 5109 SAS Hall, Campus Box 8203, Raleigh, NC 27695, U.S.A. Author Bitzer is with Department of Computer Science, College of Engineering, NC State University, 890 Oval Drive, Campus Box 8206, Raleigh, NC 27695, U.S.A. Direct Inquiries to author Breidt (E-mail: fred.breidt@usda.gov).*

doi: 10.1111/1750-3841.15091

Journal of Food Science A Publication of the Institute of Food Technologists

form of the acids as defined by the Henderson–Hasselbalch equation (reviewed in Bugler & Cogley, 1998). The pH of acidic foods and beverages is influenced by the complex and often unknown buffer components, so it is critical to be able to predict pH in a complex buffer system.

Buffer capacity (BC) models have been developed using equations for the ionic equilibria of weak acids and bases in solution (Butler & Cogley, 1998). For relatively simple defined solutions of weak acids or bases, buffering may be modeled analytically to determine pH, the concentrations of all the ions in solution, as well as uncharged acid or base species. For undefined systems with multiple unknown buffers, the BC can be estimated using the derivative of titration curves for environmental applications such as landfill leachate (Gibs, Shoenberger, & Suffet, 1982), water quality (Van Vooren, Van De Steene, Ottoy, & Vanrolleghem, 2001), and beverages including beer and wine (Dartiguenave, Jeandet, & Maujean, 2000; Li, Liu, Kang, & Zheng, 2015). Estimating the p*K* and concentration of buffers in solution requires numerical methods for nonlinear curve fitting. Existing computer or mathematical models (Gordon, 1982; Johansson & Johansson, 1979; Simms, 1926; Van De Steene, Van Vooren, Ottoy, & Vanrolleghem, 2002) have shown the utility of using the simplifying assumption that a series of monoprotic acids can represent the buffering of a complex system with unknown polyprotic or monoprotic weak acids and bases. Previous studies, however, did not combine BC models with pH prediction for food ingredients with unknown buffering, such as spices, vegetable material, and other low-acid ingredients. We have developed a combined BC–pH prediction model to aid in determining the safety and pH stability of acid and acidified foods.

For applications such as the addition of low acid food ingredients to an acid food, predicting how pH may change is important. Our objective was to determine the buffering components (as known or hypothetical monoprotic acids or bases) in acid and low acid food ingredients. Once a matrix of concentration and p*K* values has been estimated for food ingredients, the model may be used to predict pH of solutions containing those ingredients at various concentrations with other ingredients for which BC is known. We include data on pH prediction with sodium and calcium (monovalent and divalent cations) chloride, under a variety of pH conditions, to be relevant to food matrices, including recently developed calcium-based vegetable fermentations (MccFeeters & Perez-Diaz, 2010). A companion paper (Longtin, Price, Mishra, & Breidt, 2020) reports data from BC models for a variety of acid and low-acid ingredients in salad-dressing products. Once the BC of an acid or low-acid ingredient is known, the magnitude of pH changes due to the addition of that ingredient to acid or acidified foods may be estimated. The combined BC–pH prediction model may be useful for food safety applications and product development applications by enabling the prediction of equilibrated pH in food ingredient mixtures.

2. METHODS

2.1 Modeling pH of buffer solutions

To determine pH from a mixture of monoprotic weak acids and bases, the molar concentrations C_a and C_b for the acids and bases and the corresponding equilibrium constants K_a or K_b values are required. An algebraic equation may be derived from the combined weak acid equilibrium and charge balance equations to solve for $[H^+]$ (and therefore pH) with K_w being the equilibrium constant of water (10^{-14}), as described by Butler and Cogley (1998):

Table 1–Matlab program dependencies for the model.

Program name	Description
Data2Beta.m	Process titrator data to generate BC curve
fillgaps.m	Fill gaps in BC curves with a given increment and linear model
SCBCfit.m	Implement the trigonometric regression algorithm
SimplexBCPK_DF.m	Fit the buffer model to the trigonometric regression curve
Beta2Conc.m	Convert a BC value to a molar concentration for a buffer
BetaModel_AB.m	Calculate a BC curve from a matrix of buffer concentration and p*K* value
CalcpH_AB.m	Calculate pH from a matrix of buffer concentration and p*K* values
ConbineABs.m	Merge buffers with closely related p*K* values (see pK_tol, Table S3)
Conc2Beta.m	Convert a buffer concentration to a BC value
PlotBCfit.m	Plotting function to generate BC curve plots
AdjpKa_AB.m	Adjust a p*K* value for temperature and ionic strength
GetAdjC.m	Calculate the salt of an acid or base needed to optimize pH (M/L)
fmincon.m	Matlab function (optimization toolbox)
fminsearch.m	Matlab function (optimization toolbox)

$$\begin{matrix}\text{Acid term,} & \text{Base term,} & \text{Hydroxyl } [OH^-] \text{ term}\end{matrix}$$
$$[H^+] = (C_a K_a/(K_a+[H^+])) - (C_b[H^+]/([H^+]+K_b)) + K_w/[H^+] \quad (1)$$

To solve Eq. 1 for $[H^+]$, the solution of a fourth-order polynomial is required. Because the order of the polynomial equation increases with additional acid or base terms, numerical solutions may be used for estimating pH. Generalizing for multiple monoprotic acids and bases as well as rearranging and adding terms for the concentration of the salts of acids or bases A_{ni}, (a positive value for anions) or C_{ti} (negative value for cations) results in

$$0 = \sum(C_{ai}K_{ai}/(K_{ai}+[H^+])) - \sum(C_{bi}[H^+]/([H^+]+K_{bi})) + K_w/[H^+] - [H^+] + \sum(A_{ni}+C_{ti}) \quad (2)$$

An algorithm was developed using Matlab (CalcpH_AB.m; Table 1) to obtain a numerical solution to Eq. 2 for a given set of concentration, p*K* and A_n or C_t values using Newton's minimization method as suggested by Butler and Cogley (1998). A defined tolerance (10^{-10} for the error-function value, or 1,000 iterations) was used as the stopping criterion for the minimization algorithm. The concentrations of acid and base solutions used for pH validation studies were based on a Latin hypercube method (Garud, Karimi, & Kraft, 2017; McKay, Beckman, & Conover, 1979) to represent an unbiased distribution of concentrations. Except as noted, all chemicals were obtained from Sigma–Aldrich (St. Louis, MO, USA). Solutions without added salt included 0 to 50 mM acetic acid or citric acid and 0 to 10 mM ammonia and were prepared as shown (Figure 1). Similarly, acetic acid (0 to 100 mM) and ammonia solutions (0 to 20 mM; Figure 2A) were prepared containing salt mixtures of calcium chloride ($CaCl_2$) and sodium chloride (NaCl) (Figure 2B). For pH estimation, the equilibrium constants for acids and bases (transformed by the negative $\log_{10}$ to p*K* values) were adjusted as needed for ionic strength due to added sodium calcium salts using the Davies

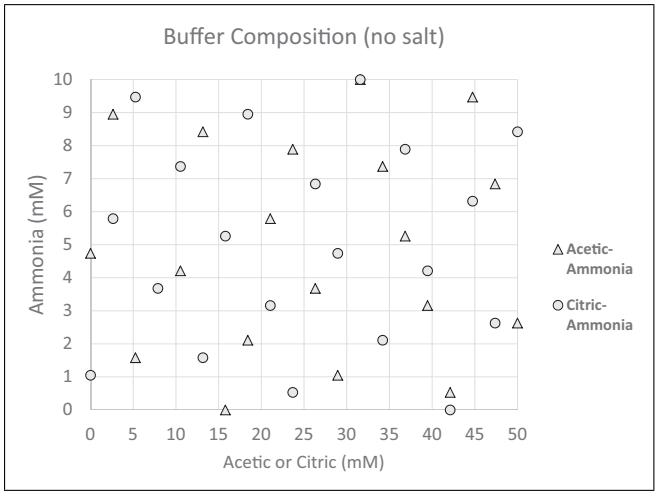

Figure 1–Buffer composition for pH validation with no added salt. The solution compositions for mixtures of acetic acid (triangles) or citric acid (circles) in combination with ammonia are shown.

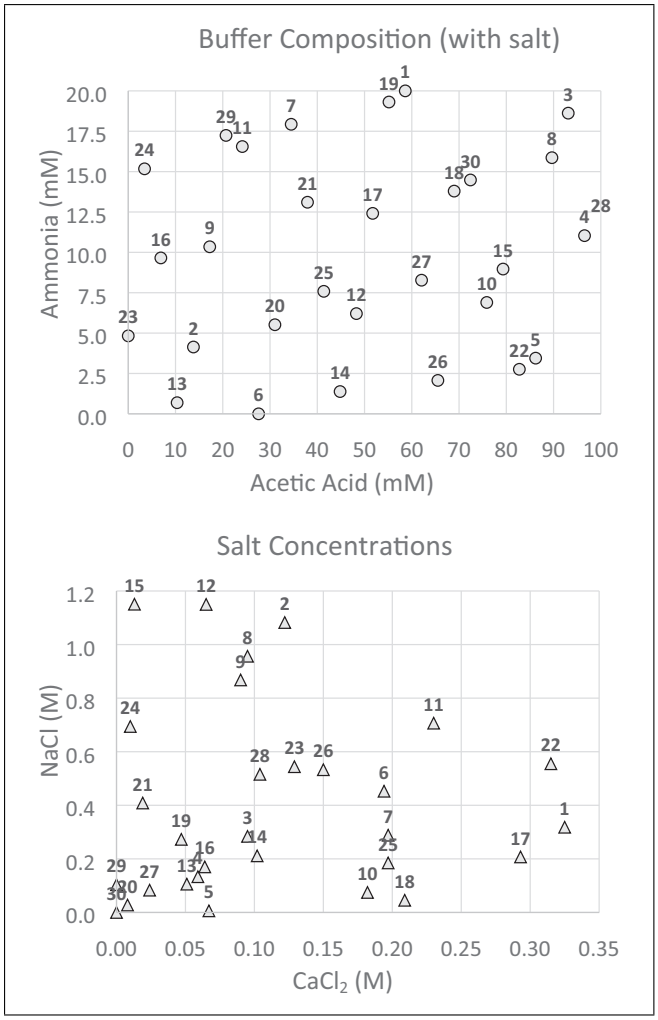

Figure 2–Buffer composition for pH validation with varying ionic strength using randomly selected combinations of NaCl (0 to 1.2 M) and $CaCl_2$ (0 to 0.325 M). Solutions containing mixtures of acetic acid and ammonia as indicated by the numbered circles (A). The corresponding numbered data points showing NaCl and $CaCl_2$ concentrations as indicated by the triangles (B).

Table 2–Example parameter file.

Parameter	Value	Description
minDeltapH	0.02	Smallest pH change allowed due to an addition of acid or base
Order	15	Number of terms in the Trigonometric regression
MinConc	0.001	Minimum buffer concentration from model predictions (M/L)
WaterSalt	2	NaCl percent for calculating the water BC curve
NaClpercent	2	NaCl percent for calculating the BC curve
NpKs	7	Initial number of p*K* values for the BC model fit
UB	12	Upper bound for the pH range for BC curve fit
LB	2	Lower bound for the pH range for BC curve fit
pK_tol	0.2	Minimum pH difference in adjacent p*K* values
trim_beg	0	Number of data points to trim from the start of the BC curve
trim_end	10	Number of data points to trim from the end of the BC curve
adjC	0	Concentration (M/L) for salt of an acid or base (cations = –val)
minGap	0.3	Minimum gap in pH for linear auto-filling BC curve
increment	0.1	Increment of pH value for gap filling of BC curve

equation (AdjpK_AB.m; Table 1) as suggested by Butler and Cogley (1998, p. 94). Further details of the concentrations used for validation studies are shown in Table S1 and S2.

2.2 Titrations

A 0.6 M solution of acetic acid with 2% NaCl was prepared from a commercial vinegar stock solution (approximately 30% acetic acid) in deionized water and 50 mL was used for titration with 1.2 N hydrochloric acid (HCl) or 1.3 N sodium hydroxide (NaOH) using an automated titrator (Model 902, Hanna Instruments, Smithfield, RI, USA). Similarly, 12.4 mM citric acid with 2% NaCl (Sigma–Aldrich) solution was prepared in water and 50 mL used for acid or base titration with 0.122 N HCl or 1.38 N NaOH. The titrator was set to use dynamic dosing over a pH range of pH 2 to 12 from the starting pH of the solution. Data were exported from the titrator as a text file and processed using a custom Python script (GetCurve.py, F. Breidt, unpublished) to generate a comma delimited data file containing two columns: the volume of acid or base added and the resulting pH. The data were then imported into Matlab with the built in csvread.m function and saved as a Matlab workspace variable. Additional workspace variables were added, including the concentration of the acid and base used for titration and the initial volume for the titration.

Matlab functions for the BC modeling procedures described below are listed in Table 1. Matlab version R2018a was used for processing the data. A single Matlab "Live-Script" file was used to sequentially call the functions and process the data. Model parameters are listed in Table 2. Parameters were stored in a text file and were imported as workspace variables at the start of the live-script. To generate a BC curve from the titration data, an iterative step-wise derivative was calculated (Butler & Cogley, 1998, p. 133) with the acid or base concentration used and volume titrated (function Data2Beta.m; Table 1):

$$\beta = \Delta\,(Acid \text{ or } Base)/\Delta\,pH \tag{3}$$

where $\Delta Acid$ or $\Delta Base$ represented the change in normality of acid or base in the solution being titrated (based on the volume added at each titration step and the concentration of the HCl or NaOH) with the resulting pH change ΔpH. A minimum change (typically 0.02 pH units) in pH was defined to prevent deviations

in the derivative if an air bubble was in the acid or base liquid delivery line of the titrator, or some other error occurred during titration (minDeltapH in Table 2). Prior to further processing, BC curves were manually trimmed by setting trim_beg or trim_end (Table 2) as needed, so the ends of each BC curve, typically at pH 2 and pH 12, had similar BC values. In addition, gaps in the BC greater than a set pH value, typically 0.3 pH units or greater (minGap; Table 2), were automatically filled with BC data points at a specified interval of 0.1 pH units (increment; Table 2) using a linear function between the end points of the gap (fillgap.m; Table 1). These steps were necessary to facilitate the subsequent use of curve fitting algorithms.

2.3 Modeling BC curves

Curve fitting for BC data was accomplished using a trigonometric least squares regression method (Eubank, 1990; Newbery, 1970), based on the equation:

$$F(x) = B_o + A_1\sin(\alpha x) + B_1\cos(\alpha x) + A_2\sin(2\alpha x) + B_2\cos(2\alpha x) + A_3\sin(3\alpha x)\ldots \quad (4)$$

where $F(x)$ represented the BC value for a given pH (represented as the variable x), A_i and B_i were scalar parameters, and α was a multiplier (0.5) used for x to fit pH values between pH 2 and 12. The model results were fit to the BC data by minimizing the error sum of squares:

$$SS(E) = \sum (F(x) - Y(x))^2 \quad (5)$$

where $F(x)$ was the trigonometric series model (Eq. 4) and $Y(x)$ was the BC curve generated from Eq. 3. To estimate parameters for the trigonometric model, a square matrix (**M**) of partial derivatives with respect to each parameter (P_i) of Equation 5 was generated:

$$\partial SS(E)/\partial(P_i) = 0 \quad (6)$$

An optimized parameter vector ($\mathbf{P_o} = B_0, A_1, B_1, A_2, B_2, \ldots$) for the terms in Eq. 4 was then obtained using Matlab's "mldivide" operator to solve the matrix equation $\mathbf{M^*P_o} = \mathbf{Y_i}$, where the vector $\mathbf{Y_i}$ consisted of the corresponding $Y(x)$ term for each partial derivative equation (implemented as SCBCfit.m; Table 1). Typically, 15 A_i and B_i parameters (defined as the parameter "Order"; Table 2) were used for the solution. The resulting trigonometric series model with optimized parameters ($\mathbf{P_o}$) represented a mathematically defined smoothed curve representing the BC data and was used as a template for modeling the BC curve with the BC model as described by Butler and Cogley (1998):

$$\beta = 2.303 \times \left(\sum \left(C_i K_i [H^+]/\left([H^+] + K_i\right)^2\right) + K_w/[H^+] + [H^+]\right) \quad (7)$$

where C_i and K_i are the concentration and dissociation constants for weak acids or bases that contribute to buffering in the solution that was originally titrated and K_w is the equilibrium constant for water.

A matrix of C_i and K_i parameter values was obtained for the optimized fit of Eq. 7 to the trigonometric function, using the constrained nonlinear minimization algorithm fmincon.m (Optimization Toolbox, Matlab) to minimize squared error values (implemented as SimplexBCPK_DF.m; Table 1). Starting values for the minimization algorithm were chosen by using a predefined number of pK values, typically 7 (parameter NpKs; Table 2), evenly distributed between the minimum and maximum pH values for the titration curve with the corresponding concentrations estimated (assuming pH = pK) from BC value from the trigonometric regression model. The NpKs parameter value could also be adjusted manually based on the complexity of the BC curve. Concentrations were constrained to positive values only, and K_i values (transformed to pK values) were constrained between the upper and lower pH values for the BC curve, typically pH 12 to 2 (defined as UB and LB, respectively, in the parameter file Table 2). For the resulting concentration–pK matrix, all pK values that were within a set tolerance of each other, typically 0.2 pH units (pK_tol; Table 2), were combined and the corresponding concentration values (C_i values) were added. The result was an N × 2 matrix of concentration and pK values for each ingredient. The estimated pH from the matrix was then determined using Eq. 2 with A_n and C_t equal to zero, with corrections of the pK due to the NaCl concentration (NaClpercent; Table 2). An additional parameter (WaterSalt; Table 2) was used for correcting the BC of water for the final BC graph. By default, pK values for buffers identified by the model that were greater than pH 7 were considered bases for pH estimation (by Eq. 2), and pK values for buffers with a pK less than or equal to pH 7 were considered acids.

To model salts of acids or bases for pH prediction from the BC model data, a positive or negative value representing the sum of A_n and C_t (add for anion or subtract for cation) was used to minimize the squared difference between the observed and predicted pH using Eq. 2 (GetAdjC.m; Table 1). The magnitude of this value was essentially the millimolar concentration of the salt of the acid or base. If no acid or base salts were present, the magnitude of this "ion" value represented the error in the model for the predicted BC curve. If known acid or base salts were present, this value could be set in the parameter file (parameter adjC; Table 2).

2.4 Model validation for pH prediction

Space-filling Latin hypercube experimental designs (McKay et al., 1979) were used for pH model validation. Simple mixtures of 0 to 50 mM citric acid or acetic acid and 0 to 10 mM ammonia were used at the concentrations indicated in Figure 1 and Table S1. A third validation solution set ($n = 30$) was generated for mixtures of 0 to 100 mM acetic acid and 0 to 20 mM ammonia that varied in ionic strength from 0 to 1.5 M using randomly selected combinations of 0 to 1.15 M sodium chloride and 0 to 0.33 M calcium chloride that cover the range that is commonly used in acidified foods (Figure 2; Table S2). Analytical grade reagents were used for all pH validation assays, including glacial acetic acid, ammonium chloride, $CaCl_2$, and NaCl (Sigma–Aldrich). Stock solutions of glacial acetic acid (0.3 M) in water, ammonium chloride (0.06 M) in water, $CaCl_2$ (1.02 M) in water, and NaCl (3.04 M) in water were prepared in 1.0-L volumetric flasks measured by w/v%. The stock solutions were stored at room temperature. Aliquots of each stock solution were added as appropriate to generate the solutions with or without added salts (as described above) in a 50 mL final volume in a volumetric flask with deionized water. Each solution was mixed by inversion before being transferred to 100 mL plastic beakers for pH measurement. The pH of each sample was measured at ambient temperature using an Accumet AR25 pH meter (Fisher Scientific) equipped with a gel-filled electrode (model FC210, Hanna Instruments), which was calibrated with certified standards of pH 2.00, 4.00, 7.00, and 10.00.

2.5 High-performance liquid chromatography analysis

High-performance liquid chromatography (HPLC) of acetic acid solutions was conducted using a modification of the method of McFeeters and Barish (2003). Separation was performed using an Aminex HPX-87H column (300 mm × 7.8 mm, Bio-Rad Laboratories, Hercules, CA, USA) with a Shimadzu HPLC system (UFLC Shimadzu Scientific Instruments, Durham, NC, USA) and the accompanying software from the manufacturer. The mobile phase was 0.01 N sulfuric acid with a flow rate of 0.9 mL/m. The column temperature was 65 °C. Acetic acid was quantified using a refractive index detector (RID-10A, Shimadzu) with a calibration curve in the range of 0.5 to 100 mM.

2.6 Data analysis

To analyze the pH predictions, a statistical model was fit to the $n = 69$ observed pH levels with the general linear models procedure of SAS (SAS Institute, Inc., Cary, NC, USA) statistical analysis software. For the observed and predicted pH for aqueous solutions with ammonia, acetic acid, and citric acid, three types of salt conditions were used: no added salts with ammonia and acetic acid (NS_A); no added salts with ammonia and citric acid (NS_C); and ammonia and acetic acid with combinations of $CaCl_2$ and NaCl (WS) (see Figure 1 and 2). The dependence of the observed pH on the prediction was allowed to vary across the three validation studies with a different slope parameter for each. The linear model was stated formally as

$$pH_i = \beta_1 X_{i1} \widehat{pH_i} + \beta_2 X_{i2} \widehat{pH_i} + \beta_3 X_{i3} \widehat{pH_i} + \epsilon_i \quad (8)$$

Here, i was an index for each observation, β_1, β_2, and β_3 denoted the three salt-specific slopes (NA_S, NA_C, and WS, respectively), pH_i was the prediction based on the Eq. 2 above, and ε_i represented the experimental error. For $j = 1$, 2, and 3, X_{ij} was an indicator variable taking the value of 1 if observation i comes from salt level j, and otherwise was 0. For analysis of buffer concentrations from titration data, a second-order polynomial regression for measured and predicted buffer concentrations was done in Excel. The Matlab programs and related files are available at (https://www.ars.usda.gov/southeast-area/raleigh-nc/food-science-research/).

3. RESULTS AND DISCUSSION

To analyze the quality of pH predictions, a statistical model was fit to the $n = 69$ observed pH levels for buffer solutions containing ammonium chloride, acetic acid, or citric acid with and without mixtures of $CaCl_2$ and NaCl (Figure 1 and 2). Solution pH was measured and compared to model predictions (Figure 3). The majority of pH values were targeted to be between pH 3 and pH 4 to be relevant for most acid and acidified food products. The model algorithm (CalcpH_AB.m; Table 1) provided an excellent fit for describing the association between the observed and predicted pH, with a coefficient of determination (r^2) = 0.9915. The root mean squared error from this model was 0.138. With an average observed pH in the study of 3.35, this results in a coefficient of variation that was less than 4%. Although all estimated slopes were close to 1, there was a significant difference among them ($F = 10.7$, $P < 0.0001$, $df = 2, 66$). In particular, the slopes for both no-salt treatments (NS_A and NS_C) were significantly lower than the slope for the salt treatment (WS; Table 3). Figure 3A shows the overlays of the general linear statistical model with the regression lines corresponding to the three salt treatments. Figure 3B shows that the relative error of prediction stays roughly constant over a wide range of pH values. A weakness is in evidence for the predictions for NS_A, as they are all less than their corresponding observed values, with negative relative error. However, the magnitudes of the errors were within 10% of the observed pH values.

For solutions containing mixtures of acid or acidified foods, which may contain low-acid food ingredients, pH prediction is problematic because the buffers of the system may be undefined.

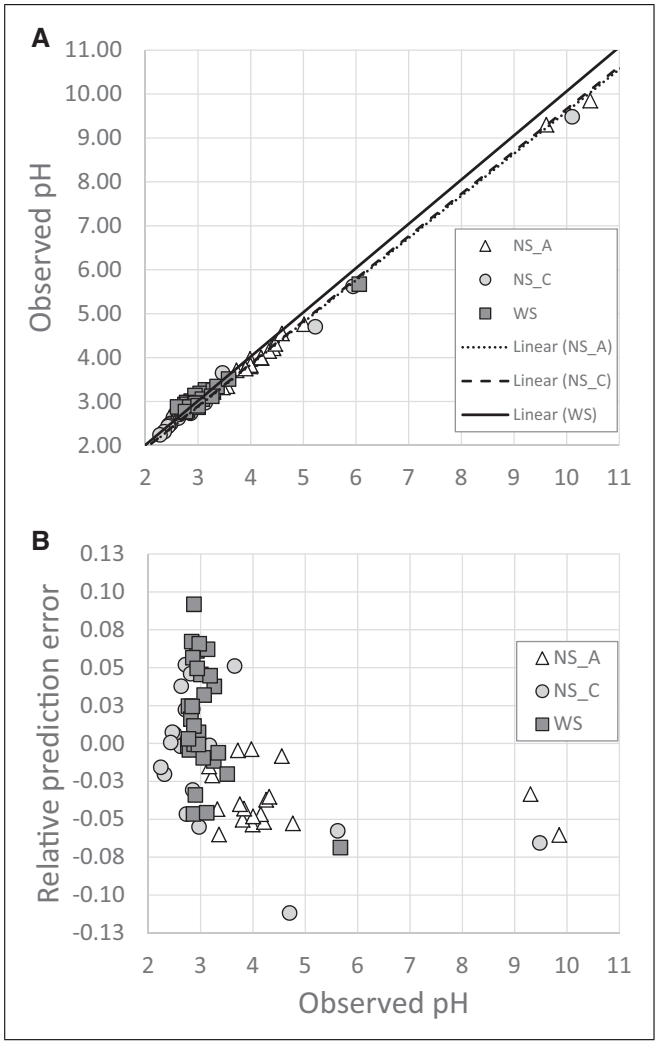

Figure 3–Observed and predicted pH data. The relationship between the observed (measured) pH values and the predicted pH values from the model (panel A) and the relative error for each data point (panel B). The buffer solutions consisted of acetic acid and ammonia solutions without added salts (NS_A, triangles); citric acid and ammonia solutions without added salts (NS_C, circles); and acetic acid and ammonia solutions with NaCl and/or $CaCl_2$ (WS, squares). The lines in panel A represent linear regression for each data set: NS_A, dotted line; NS_C, dashed line; WS, solid line.

Table 3–Slope parameters for regression analysis.

Slope parameter	Salt and acid type[a]	Least squares estimate	Standard error	95% Confidence limits
β_1	NS_A	0.9653	0.0071	(0.9478, 0.9828)
β_2	NS_C	0.9600	0.0069	(0.9431, 0.9769)
β_3	WS	1.0063	0.0081	(0.9864, 0.10263)

[a]Salt types: NS_A, no salt acetic acid and ammonia; NS_C, no salt citric and ammonia; WS, acetic acid and ammonia treatments with $CaCl_2$ and NaCl mixtures.

However, a reasonable assumption is that a system of monoprotic buffers may be used to represent the complex buffering in these food mixtures if the concentration and pK values can be estimated (Simms, 1926; Gordon, 1982). By removing the complexity of modeling polyprotic acids, the task of estimating pH from complex buffer systems became simpler, and the pH estimation algorithm described above (Eq. 2) for monoprotic acids and bases was therefore used.

To determine the BC of acid or acidified food products and use this to predict pH changes in the foods as additional ingredients may be added, we developed a BC model using a combination of trigonometric regression and nonlinear curve fitting methods. The algorithm consisted of a Matlab "Live-Script" file with an associated parameter file and Matlab functions (Table 1 and 2). The results illustrating the BC modeling process and pH prediction for a 0.6 M acetic acid solution with 2% NaCl are shown in Figure 4. Titration data were first converted to BC using Eq. 3 as implemented in the Data2Beta.m algorithm (Table 1; Figure 4A). To facilitate curve fitting, we found that it was helpful to have BC values at the ends of the BC curve to be approximately equal, so the data were manually trimmed (Figure 4A, blue circles) by setting the trim_beg and trim_end values in the parameter file (Table 2). Gaps in the titration curve were automatically filled with a linear model between gap end points based on parameters set in a parameter file (Table 2).

For the trigonometric regression (SCBCfit.m algorithm; Table 1; Eq. 4), we found that $\alpha = 0.5$ resulted in a BC model for pH 2 to 12 that had model results similar to observed BC data. By default, we used 15 sine–cosine paired terms, which were sufficient to approximate most BC curves generated for acid and low-acid food ingredients (Longtin et al., 2020). The trigonometric regression results for the acetic acid BC data are shown in Figure 4B (black line). Subsequently, the BC model (Eq. 7) was used to fit the smoothed trigonometric regression line representing the BC data with the Matlab constrained nonlinear curve-fitting algorithm, fmincon.m (Table 1; Figure 4C). For the acetic acid data, the pK values were constrained by the upper (UB) and lower (LB) pH bounds, typically defined as pH 12 and 2, respectively, in the parameter file (Table 2). The initial pK and concentration values for the curve fitting algorithm included seven (NpKs = 7; Table 2) evenly distributed pH values between pH 2 and 12. Similar pK values were combined during processing based on a defined tolerance (pK_tol = 0.2; Table 2) to reduce the number of pK values to approximate the BC curve. For the acetic acid solution, a single pK was estimated by the model, at a pH value of 4.52. This pH was approximately equal to the expected pK of 4.49 for a 0.6 M acetic acid solution in 2% NaCl (ionic strength of 0.342). The pK correction for ionic strength was done using the Davies equation (Butler & Cogley, 1998, p. 49).

To estimate a value for the salt of an acid or base needed to correct the predicted pH based on measured and estimated pH values from the pK-concentration matrix, the Matlab fminsearch.m algorithm was used to estimate an A_n (positive) or C_t (negative) value from Eq. 2. For titration of known salts of an acid or base, this A_n or C_t value may be set as a positive or negative value in the parameter file (adjC; Table 2) if it is known that the salt of an acid was used for titration. For the acetic acid solution used, AdjC was initially set to zero. The unadjusted pH estimated by the estimated pK and buffer concentration was 2.31 and the measured pH (mean of the initial pH for the acid and base titrations) was 2.41. This required a cation correction of 2.35 mM (modeled as $C_t = -2.35$ in Eq. 2) for the estimated pH to match the observed pH. It is possible that the error in the model prediction for pH was incorrect due to other factors than a cation salt of the acid, so this correction should only be considered one possible explanation for the error, which could be due to numerical inaccuracies or errors in titration. In addition to the estimated pK value, the

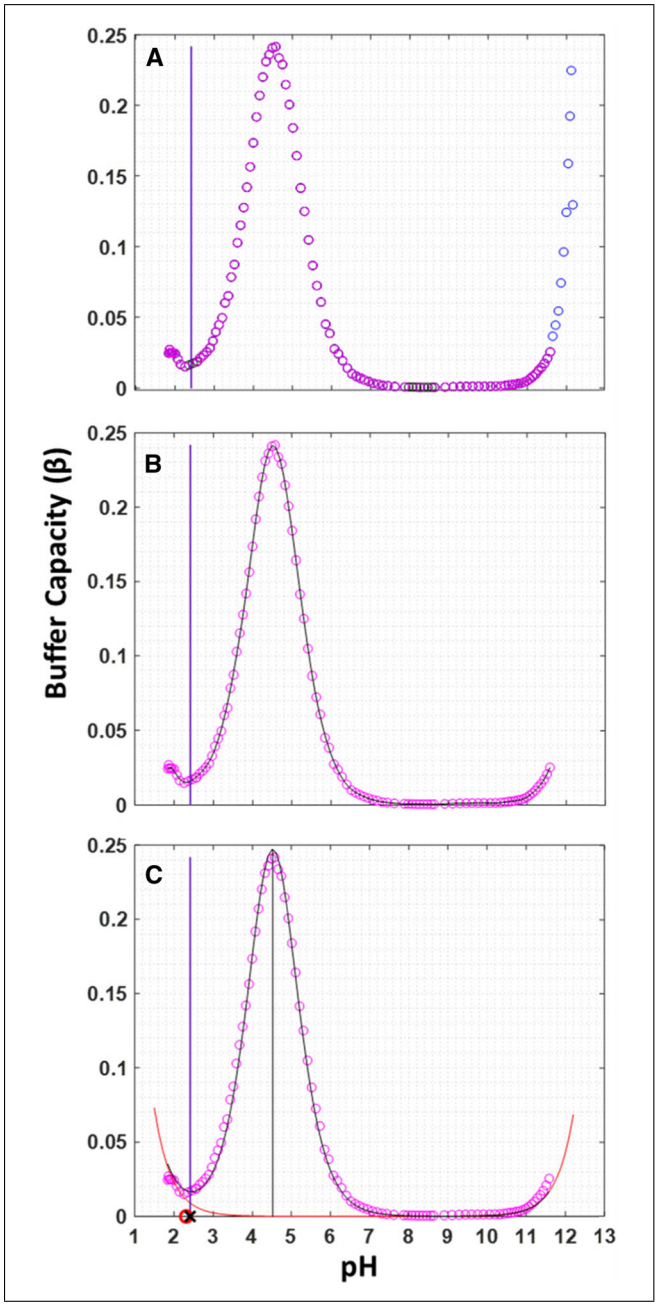

Figure 4–The BC modeling method for 0.6 M acetic acid. A commercial vinegar solution (approximately 0.6 M acetic acid) was titrated with 1.22 N HCl and then with 1.3 N NaOH. The BC curve (purple circles) that is the derivative of the combined titration curves (A) was trimmed (blue circles) to result in a BC curve with symmetrical ends, followed by linear gap filling (black circles). The vertical red and blue lines represent the initial pH of the acid and base titrations, respectively. The trigonometric regression (B, black curve) was fit to the BC data line. Panel C shows the BC model fit (black curve), the BC of water (red curve), the pH calculated from the mean of the initial titration values (black x), and the predicted pH from the BC model unadjusted for C_t (red circle); the pK (pH value) and BC (β) value for the predicted buffer are represented by the black vertical line.

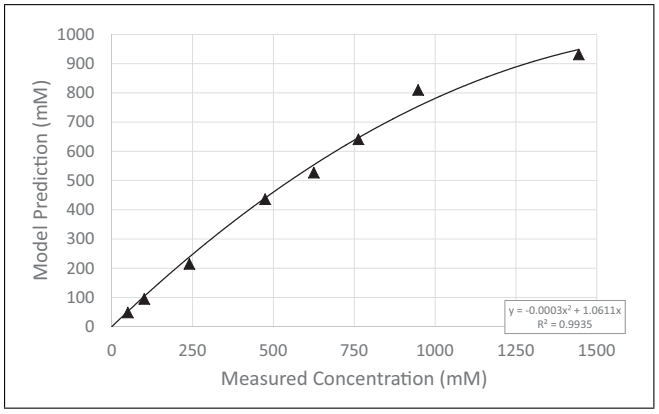

Figure 5–Measured and predicted acetic acid concentration. The concentration of the acetic acid buffer in solution as measured by HPLC (*X* axis) and as predicted by the BC model (*Y* axis) is shown (triangles). The line represents a second-order polynomial regression, with an $r^2 = 0.9935$.

Table 4–Buffers predicted for the citric acid solution.

p*K*[a]	Adjusted p*K*[b]	Predicted p*K*[c]	Concentration (mM)	Buffer capacity
3.13	2.86	3.01	12.50	0.0094
4.76	4.49	4.51	13.36	0.0078
6.39	6.12	5.74	10.98	0.0063

[a]p*K* value from CRC (1995)
[b]p*K* value adjusted for 2% NaCl using the Davies equation
[c]p*K* value predicted by the BC model

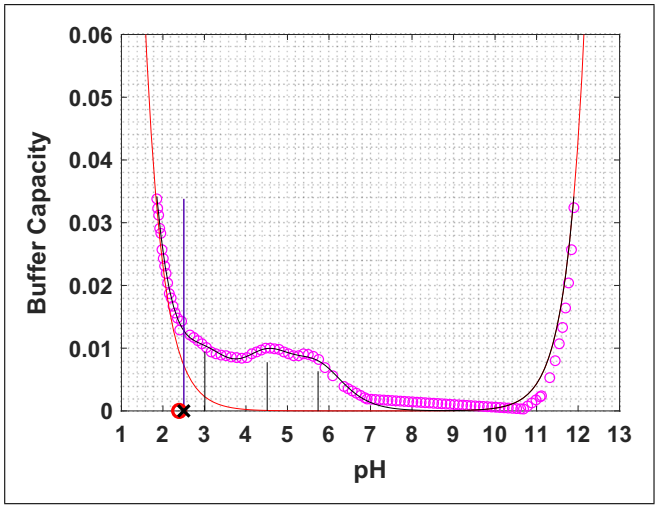

Figure 6–The BC model of 12.4-mM citric acid. The three monoprotic buffers identified by the model representing citric acid had p*K* values of 3.01, 4.51, and 5.74 (black vertical lines). The observed and predicted pH (black x and red circle on the *x*-axis) was 2.50 and 2.39, respectively.

pH prediction also depended on the estimated concentration of the acetic acid derived from the BC model.

Unexpectedly, the predicted buffer concentration for the acetic acid solution based on the model prediction was 429.3 mM, but the measured acetic acid concentration by HPLC was 636.0 mM. To investigate this difference, we examined a series of concentrations for acetic acid solutions without added NaCl (triangles; Figure 5). The data were found to fit a polynomial model with a regression coefficient (r^2) value of 0.99. For values less than 400 mM, the measured and predicted values were within 6%; however, for higher concentrations the error increased to approximately 10% at 500 mM and approximately 24% for 1 M. The reason for this deviation in estimated buffer concentration from the measured values at high concentrations (above 400 mM) remains unclear, although it may be related to ion activity coefficients, which are influenced by ionic strength (Butler & Cogley, 1998, p. 49). This will be the subject of future investigation. Despite this difference, however, the estimated pH from the predicted p*K* of 4.52 and concentration of 429 mM acetic acid was only 0.1 pH units different from the observed value (2.31 predicted compared with 2.41 observed). However, this approximated the pH predicted for the acetic acid concentration of the titrated solution as measured by HPLC (636 mM) which was 2.36.

A solution of citric acid in 2% NaCl was titrated and the resulting matrix of monoprotic buffers was used to calculate solution pH. The BC model predicted three monoprotic acid buffers (Figure 6; Table 4). The mean of the initial pH from the acid and base titrations was pH 2.50 ± 0.003, and the predicted pH from the BC model using Eq. 2 was 2.39. The three estimated concentrations for predicted buffers (Table 4) had a mean value of 12.28 ± 1.209 mM, closely approximating the 12.4 mM citric acid solution used for titration. The p*K* values for the predicted buffers were similar to the p*K* values for citric acid adjusted for 2% NaCl, with the predicted 3.01, 4.51, and 5.74 p*K* buffers differing from the calculated p*K* values (adjusted for 2% NaCl) by –0.15, –.002, and 0.38 pH units, respectively. The reason for these differences may be due to titration error or other unknown factors; however, the predicted pH error from these model buffers was approximately 0.1 pH units, similar to errors obtained in the pH validation results shown in Figure 3. These results support the assumption that polyprotic acids may be modeled as a composite of monoprotic acid buffers as proposed by Simms (1926) and Gordon (1982).

4. CONCLUSIONS

Models for estimating pH, p*K*, and concentrations of buffers in solution were developed and implemented in a Matlab "live-script" format. By combining buffer modeling with pH prediction, the pH of solutions with complex buffering could be estimated. The methods were based on standard ionic equilibria equations and were implemented with a two-step modeling procedure, involving a trigonometric regression, followed by a nonlinear curve-fitting algorithm to estimate concentration and p*K* values for buffer solutions. Concentrations were underestimated for acetic acid greater than 400 mM, so further research may be needed. The models may be useful for regulatory purposes to help predict pH changes and pH stability of acid food products such as salad dressings that have small amounts of low-acid ingredients. The models may also be useful for product development to predict the pH of mixtures of acid and low-acid food ingredients.

ACKNOWLEDGMENTS

The authors thank Mrs. Sandra Parker for assistance with manuscript preparation. This research was funded in part (agreement 58-6070-8-0019F) by the Association of Dressings and Sauces, Atlanta, GA, as well as by Pickle Packers International, Inc., Washington, DC, and the USDA-ARS Food Science Research Unit, Raleigh, NC.

AUTHOR CONTRIBUTIONS

Author Price led the laboratory efforts with authors Longtin and Conley-Payton for validation of the pH and titration data. Authors Osborne and Johanningsmeier contributed to the statistical analysis and experimental design for validation experiments,

respectively. Authors Bitzer and Breidt developed the modeling concepts and author Breidt wrote the Matlab algorithms. All authors contributed to writing the manuscript and preparation of the figures and tables.

REFERENCES

Breidt, F., Jr. (2006). Safety of minimally processed, acidified and fermented vegetable products. In G. M. Sapers, J. R. Gorny, & A. E. Yousef (Eds.), *Microbiology of fruits and vegetables* (pp. 313–335). Boca Raton, FL: CRC Press.

Butler, J. N., & Cogley, D. R. (1998). *Ionic equilibrium: Solubility and pH calculations*. New York, NY: John Wiley and Sons.

CRC. (1995). *Handbook of chemistry and physics* (76th ed.). New York, NY: CRC Press.

Dartiguenave, C., Jeandet, P., & Maujean, A. (2000). Study of the contributions of the major organic acids of wined to the buffering capacity of wine in model solutions. *American Journal of Enology and Viticulture*, *51*(4), 352–356.

Eubank, R. L., & Speckman, P. (1990). Curve fitting by polynomial-trigonomietric regression. *Biometrika*, *77*(1), 1–9.

FDA. (1979). *Acidified foods. 21 CFR 114*. Washington, DC: U.S. Food and Drug Administration.

Garud, S. S., Karimi, I. A., & Kraft, M. (2017). Design of computer experiments: A review. *Computers and Chemical Engineering*, *106*, 71–95.

Gibs, J., Shoenberger, R. J., & Suffet, I. H. (1982). A simplified buffer capacity model for sanitary landfill lechate. *Water Research*, *16*, 699–705.

Gordon, W. E. (1982). Data analysis for acid-base titration of an unknown solution. *Analytical Chemistry*, *54*, 1595–1601.

Hosein, A. M., Breidt, F., & Smith, C. E. (2011). Modeling the effects of sodium chloride, acetic acid, and intracellular pH on survival of *Escherichia coli* O157:H7. *Applied and Environmental Microbiology*, *77*(3), 889–895.

Ito, K. A., & Chen, J. K. (1978). Effect of pH on growth of *Clostridium botulinum* in foods. *Food Technology*, *32*, 71.

Johansson, A., & Johansson, S. (1979). General-purpose program for evaluating potentiometric acid-base titrations. *Analyst*, *104*, 601–612.

Li, H., Liu, F., Kang, L., & Zheng, M. (2015). Study on the buffering capacity of wort. *Journal of the Institute of Brewing*, *122*, 138–142.

Longtin, M., Price, R. E., Mishra, R., & Breidt, F. (2020). Modeling the buffer capacity of ingredients in salad dressing products. *Journal of Food Science*.

McFeeters, R. F., & Barish, A. O. (2003). Sulfite analysis of fruits and vegetables by high-performance liquid chromatography (HPLC) with ultraviolet spectrophotometric detection. *Journal of Agricultural and Food Chemistry*, *51*(6), 1513–1517.

McFeeters, R. F., & Perez-Diaz, I. M. (2010). Fermentation of cucumbers brined with calcium chloride instead of sodium chloride. *Journal of Food Science*, *75*(3), C291–C296.

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *42*(1), 55–61.

Neta, E. R. D., Johanningsmeier, S. D., Drake, M. A., & McFeeters, R. F. (2009). Effects of pH adjustment and sodium ions on sour taste intensity of organic acids. *Journal of Food Science*, *74*(4), S165–S169.

Newbery, A. C. R. (1970). Trigonometric interpolation and curve fitting. *Mathematics of Computation*, *24*(112), 869–876.

Russel, J. B. (1992). Another explanation for the toxicity of fermentation acids at low pH: Anion accumulation versus uncoupling. *Journal of Applied Bacteriology*, *73*, 363–370.

Simms, H. S. (1926). Dissociation of polyvalent substances I. Relation of constants to titration data. *Journal of the American Chemical Society*, *48*, 1239–1250.

Van De Steene, M., Van Vooren, L., Ottoy, J.-P., & Vanrolleghem, P. A. (2002). Automatic buffer capacity model building for advanced interpretation of titration curves. *Environmental Science & Technology*, *36*, 715–723.

Van Vooren, L., Van De Steene, M., Ottoy, J.-P., & Vanrolleghem, P. A. (2001). Automatic buffer capacity model building for the purpose of water quality monitoring. *Water Science and Technology*, *43*(7), 105–113.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Observed and predicted pH for solutions with no added salt.

Table S2. Observed and predicted pH for solutions with no added salt.