# Correlation Between pH and Composition of Foods Comprising Mixtures of Tomatoes and Low-Acid Ingredients

G. M. SAPERS, JOHN G. PHILLIPS, and ANITA M. DiVITO


## ABSTRACT

Acidity and pH data for more than 100 products comprising mixtures of tomatoes and low-acid ingredients were compared to develop a generalized method of pH prediction. Products and their major ingredients were titrated with NaOH or acetic acid to pH 4.6 and 8.1 endpoints. Two indices of acidity, developed from these data, were correlated with product pH values, and regression equations for pH prediction were obtained. The accuracy of prediction was improved by correlating the data for related products such as soups, sauces, and simple tomato-vegetable mixtures. Correlation coefficients as high as 0.9 were obtained with the last category. These results demonstrate the feasibility of pH prediction from recipe data.


## INTRODUCTION

A MAJOR CONSIDERATION in developing safe processes for home-canned foods is the product pH (Ito and Chen, 1978; Odlaug and Pflug, 1978). Values of the pH for most single-ingredient food products fall within a well-defined range (Harrel and Thelen, 1959). However, because of variability in the ingredient proportions and composition, pH values for products comprising combinations of ingredients are difficult to predict. Guidelines for processing combinations have been developed (IFT, 1977) but are of limited value with predominantly high-acid tomato-based products such as juice blends, stewed tomatoes, and meatless tomato sauces (Sapers et al., 1982a).

In a previous study (Sapers et al., 1982a), we determined the pH limits of important categories of tomato-based combinations, including many regional and ethnic foods, based on the analysis of representative products. During the course of that investigation, we evaluated several indices of ingredient acidity that might be used as predictors of product pH, obtaining high correlations between pH and acidity with several product categories. In a subsequent study (Sapers et al., 1982b), we developed a model, simulating tomato-based combinations, from which regression equations for pH prediction were derived. However, these equations only were applicable to products for which the major ingredients were among those used to develop the model, i.e., tomatoes, green bell peppers, mushrooms, carrots, red kidney beans, chopped onions, celery, chicken broth, beef broth, and ground beef.

Our objective in the present study was to develop a general method of pH prediction, based on correlations between product pH and ingredient acidity, that would be applicable to a wider range of combination products. Such a general method might be useful to Extension Specialists in screening home-canning recipes to determine the suitability of the specified thermal processing conditions, based on whether the recipes yielded high-acid or low-acid products.

## MATERIALS & METHODS

### Selection of product categories and representative products

The scope of this study was limited to categories of closely related products classified previously as high-acid foods (pH ≤ 4.6) or as borderline low-acid foods, the pH of which straddled 4.6, the boundary between low-and high-acid foods (Sapers et al., 1982a). These included various tomato soups, meatless tomato sauces, and other tomato-vegetable combinations such as juice blends and stewed tomatoes, but excluded highly acidic products such as relishes, chili sauces, and some Mexican and barbecue sauces as well as very low-acid products such as chili con carne, other tomato-bean combinations, and spaghetti sauces with meat since we foresaw no need to predict their pH. To obtain correlations between pH and acidity, we compiled data for more than 100 representative products within the applicable product categories that had been prepared and analyzed in 1979 and 1981 (Sapers et al., 1982a; DiVito et al., 1982). While only 28% of these products were intended specifically for home canning, all were considered of potential interest to home canners. Product recipes specifying ingredient proportions for these products were stored in our recipe databank.

### Indices of acidity

In addition to titratable acidity values, determined by titrating product homogenates with 0.1N NaOH to a pH 8.1 endpoint, we computed two other indices of acidity for each of the products tested based on the titration of homogenates of individual ingredients with 0.1N NaOH or 5% acetic acid (distilled white vinegar) to a pH 4.6 endpoint, as described previously (Sapers et al., 1982a). The first of these indices, the "acidity ratio," is the summation for all low-acid ingredients of the milliequivalents of acetic acid required to lower the pH of each low-acid ingredient in 100g of product to 4.6 divided by the summation for all high-acid ingredients of the milliequivalents of NaOH required to elevate the pH of each high-acid ingredient in 100g of product to pH 4.6:

$$\text{Acidity ratio} = \frac{\Sigma(\text{low-acid ingred. \%}) \times (\text{meq acetic acid/g low-acid ingred.})}{\Sigma(\text{high-acid ingred. \%}) \times (\text{meq NaOH/g high-acid ingred.})}$$

The second index, the "acidity difference," is the difference between the denominator and the numerator of the acidity ratio:

$$\text{Acidity difference} = \Sigma(\text{high-acid ingred. \%}) \times (\text{meq NaOH/g high-acid ingred.}) - \Sigma(\text{low-acid ingred. \%}) \times (\text{meq acetic acid/g low-acid ingred.})$$

Values of these indices were computed from ingredient proportions spec-

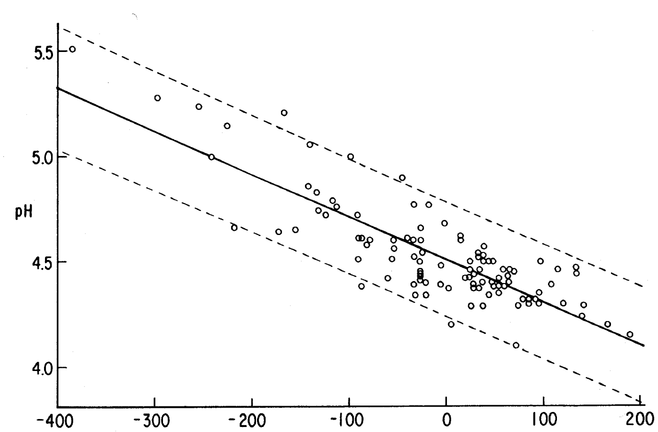


*Fig. 1—Correlation between pH and the acidity difference for tomato-based soups, meatless sauces, and other vegetable combinations (data set A + B + C).*

*Authors Sapers, Phillips, and DiVito are with the USDA-ARS Eastern Regional Research Center, 600 E. Mermaid Lane, Philadelphia, PA 19118.*

ified by the recipes and ingredient titration data obtained experimentally in duplicate, which are summarized in Tables 1-3.

### Correlations and regression analysis

Relationships between product pH and the indices of acidity for each product category were examined by means of the Statistical Analysis System (SAS), General Linear Models Procedure, yielding correlation coefficients, standard errors, and regression coefficients for equations relating pH to the indices, and comparisons of predicted vs measured pH values for each individual product. In addition, data for related product categories were pooled into three broader categories, namely, soups (A), meatless tomato sauces (B), and other tomato-vegetable combinations (C), and these data sets were further combined to give A + B, A + C, B + C, and A + B + C. The statistical analyses described above were performed on each of the pooled data sets in order to generate accurate predictive relationships having the widest applicability.

The data sets obtained for products prepared in 1979 and 1981 were analyzed separately as well as in one combined data set. Regression equations derived from the 1979 data were validated with the 1981 data by comparing predicted and measured pH values for each product.

## RESULTS & DISCUSSION

THE DEGREE OF CORRELATION between product pH and the various indices of acidity differed greatly among the individual product categories studied (Table 4). Significant correlations were obtained with tomato-mushroom sauces, creole sauces, tomato-vegetable juice blends, and stewed tomatoes. However, regression slopes varied greatly from category to category. With product categories yielding significant correlations, the pH generally was more highly correlated with the acidity ratio or difference than with the titratable acidity. This is to be expected since the acidity ratio and difference were derived from data obtained by titrating ingredients to a pH 4.6 endpoint, a value close to or within the pH range of the combination products in this study. Titratable acidity, however, is based on a pH 8.1 endpoint, a value that is far removed from the pH range of interest and likely to include buffering of little relevance to the product pH. Previously, we obtained relatively poor correlations between the pH and titratable acidity of different tomato cultivars (Sapers et al., 1977). Paulson and Stevens (1974) developed quantitative relationships explaining the contribution of different buffer systems in tomatoes to the titratable acidity and pH over a wide range of tomato pH values.

Low correlations and varying values of the regression slope, obtained with product categories other than the four mentioned above, were due largely to the relatively narrow pH ranges represented by each set of data points. Meaningful correlations were not obtained with soups and some mixed vegetable categories because of the small number of data points available.

Significantly improved results were obtained when the data for related product categories were pooled. Statistics for correlations based on the pooled 1979 data sets are given in Table 5. The best correlations (largest correlation coefficient and smallest standard error) were obtained with data set C which included juice blends, tomato purees and pastes, stewed tomatoes, and other mixed veg-

*Table 1—pH, titratable acidity, and titration values to pH 4.6 of common high-acid ingredients of tomato-based combinations*

| Ingredient | pH | Titratable acidity (pH 8.1 end-point)[a] | Titration value to pH 4.6[a] |
|---|---|---|---|
| A-1 sauce | 3.4 | 37.9 | 15.8 |
| Chili sauce | 4.0 | 19.6 | 5.4 |
| Ketchup | 3.7 | 24.2 | 8.4 |
| Lemon juice, bottled | 2.9 | 83.2 | 41.1 |
| Mustard, prepared | 3.6 | 60.1 | 23.6 |
| Olives, Spanish | 3.5 | 5.9 | 3.4 |
| Peppers, ancho chile (dried)[b] | 4.5 | 8.8 | 0 |
| Peppers, pickled | 3.3 | 23.9 | 11.9 |
| Pimientos, canned | 4.0 | 4.9 | 1.2 |
| Salad dressing, Italian | 3.3 | 27.6 | 21.9 |
| Sour cream | 4.4 | 8.6 | 1.2 |
| Tabasco sauce | 3.4 | 142.1 | 65.4 |
| Taco sauce | 4.0 | 8.4 | 2.2 |
| Tomatillos, canned | 3.8 | 19.5 | 6.2 |
| Tomato juice | 4.2 | 5.6 | 1.2 |
| Tomato paste | 4.3 | 22.5 | 4.7 |
| Tomato puree | 4.3 | 12.0 | 2.2 |
| Tomato sauce | 4.2 | 7.6 | 1.3 |
| Tomato soup, condensed | 4.2 | 5.8 | 1.1 |
| Tomatoes, canned | 4.4 | 6.4 | 1.0 |
| Vinegar | 2.7 | 86.0 | 45.0 |
| Wine, red | 3.2 | 5.5 | 3.0 |
| Wine, sherry | 3.3 | 5.4 | 2.8 |
| Wine, white | 3.1 | 6.9 | 4.1 |
| Worcestershire sauce | 3.6 | 51.4 | 21.2 |

[a] Milliequivalents of NaOH per 100g ingredient.
[b] 25g in 75 ml $H_2O$.

*Table 2—pH, titratable acidity, and titration values to pH 4.6 of common low-acid vegetable ingredients of tomato-based combinations*

| Ingredient | pH | Titratable acidity (pH 8.1 end-point)[a] | Titration value to pH 4.6[b] |
|---|---|---|---|
| Beans, green | 5.9 | 1.4 | 3.3 |
| Beans, kidney | 5.8 | 2.9 | 10.4 |
| Beans, lima | 6.2 | 2.2 | 8.7 |
| Cabbage | 5.8 | 1.1 | 2.5 |
| Carrot | 5.0 | 1.3 | 1.2 |
| Celery | 5.6 | 1.1 | 2.8 |
| Chili powder[c] | 4.6 | 59.0 | 0 |
| Corn | 7.1 | 0.4 | 5.5 |
| Escarole | 5.7 | 1.6 | 3.3 |
| Kohlrabi | 5.9 | 1.5 | 2.5 |
| Leek | 5.8 | 2.2 | 5.2 |
| Mushroom | 6.0 | 1.2 | 3.0 |
| Okra | 5.8 | 1.6 | 2.2 |
| Onion | 5.0 | 1.6 | 1.4 |
| Parsley | 5.8 | 2.0 | 7.9 |
| Peas | 6.8 | 1.8 | 9.4 |
| Pepper, bell, green | 5.0 | 1.7 | 1.3 |
| Pepper, Anaheim | 5.5 | 1.4 | 2.7 |
| Pepper, jalapeño | 5.3 | 2.0 | 2.2 |
| Pepper, serrano | 5.6 | 2.0 | 4.0 |
| Potato, white | 6.0 | 3.6 | 6.8 |
| Spinach | 6.5 | 0.8 | 4.2 |
| Turnip | 5.6 | 1.6 | 3.3 |
| Zucchini | 6.5 | 0.6 | 3.9 |

[a] Milliequivalents of NaOH per 100g ingredient.
[b] Milliequivalents of acetic acid per 100g ingredient.
[c] 5g in 50 ml boiling $H_2O$, cooled to room temperature.

*Table 3—pH, titratable acidity, and titration values to pH 4.6 of other common low-acid ingredients of tomato-based combinations*

| Ingredient | pH | Titratable acidity (pH 8.1 end-point)[a] | Titration value to pH 4.6[b] |
|---|---|---|---|
| **Animal products** | | | |
| Anchovy | 5.2 | 27.9 | 11.5 |
| Bacon, lean | 6.1 | 14.0 | 23.6 |
| Beef, ground | 6.3 | 2.9 | 6.0 |
| Bouillon cube[c] | 4.9 | 31.9 | 9.3 |
| Broth, beef | 6.0 | 0.6 | 2.1 |
| Broth, chicken | 6.3 | 1.6 | 3.7 |
| Broth, clam | 6.4 | 0.6 | 1.2 |
| Clams, minced | 6.5 | 1.9 | 10.7 |
| Ham | 5.9 | 8.6 | 15.2 |
| Pork | 6.1 | 4.1 | 7.7 |
| Sausage | 6.2 | 4.1 | 9.6 |
| Veal | 6.1 | 2.7 | 6.4 |
| **Dairy products** | | | |
| Cream, heavy | 7.0 | 1.2 | 7.7 |
| Parmesan cheese[d] | 5.6 | 14.7 | 49.2 |
| **Pasta and cereal products** | | | |
| Barley[e] | 5.9 | 0.4 | 4.2 |
| Elbow macaroni[c] | 5.8 | 2.6 | 4.0 |
| Flour[c] | 5.6 | 3.6 | 5.0 |
| Rice[e] | 6.4 | 0.7 | 1.2 |
| Soy sauce | 4.9 | 0.6 | 14.6 |

[a] Milliequivalents of NaOH per 100g ingredient.
[b] Milliequivalents of acetic acid per 100g ingredient.
[c] Reconstituted in boiling $H_2O$ in proportion of 3.8, 5.0, and 1.0g per 50 ml for bouillon cubes, elbow macaroni, and flour, respectively; then cooled to room temperature.
[d] 10g in 40 ml.
[e] Based on cooked weight.

*Table 4—Correlation between product pH and indices of acidity for selected categories of tomato-based combinations (pooled 1979 and 1981 data sets)*

| Product category | Recipes tested | Measured pH range | Titratable acidity | | Acidity ratio | | Acidity difference | |
|---|---|---|---|---|---|---|---|---|
| | | | $R^{2a}$ | Slope[a] | $R^2$ | Slope | $R^2$ | Slope |
| **Tomato sauces, meatless** | | | | | | | | |
| U.S. style | 13 | 4.24-4.66 | 0.29 | −0.049 | 0.22 | 0.108 | 0.36 | −0.0014 |
| Tomato-mushroom | 6 | 4.29-5.00 | 0.40 | −0.192 | 0.74[b] | 0.207 | 0.71[b] | −0.0031 |
| Italian spaghetti | 9 | 4.29-4.52 | 0.00 | −0.002 | 0.17 | 0.030 | 0.19 | −0.0005 |
| Marinara | 5 | 4.42-4.54 | 0.40 | −0.054 | 0.03 | −0.030 | 0.09 | −0.0005 |
| Mexican tomato-pepper | 8 | 4.29-4.77 | 0.24 | −0.032 | 0.36 | 0.098 | 0.40 | −0.0015 |
| Creole | 10 | 4.20-5.21 | 0.72[c] | −0.126 | 0.89[c] | 0.087 | 0.65[c] | −0.0028 |
| **Other tomato-vegetable combinations** | | | | | | | | |
| Vegetable juice blends | 11 | 4.20-5.06 | 0.55[b] | −0.126 | 0.89[c] | 0.155 | 0.66[c] | −0.0016 |
| Tomato puree and paste | 9 | 4.37-4.53 | 0.02 | −0.003 | 0.06 | 0.084 | 0.06 | −0.0008 |
| Stewed tomatoes | 6 | 4.15-4.54 | 0.04 | −0.016 | 0.29 | 0.454 | 0.90[c] | −0.0022 |

[a] R and slope are the correlation coefficient and slope for the regression of pH on the index of acidity, respectively.
[b] $p < 0.05$ for F-test of $R = 0$.
[c] $p < 0.01$ for F-test of $R = 0$.

*Table 5—Correlations between pH and acidity ratio or difference for pooled 1979 data sets*

| Pooled data set | Recipes tested | Acidity ratio | | | | Acidity difference | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $R^{2a}$ | Regression coeff. ± SD: Intercept | Regression coeff. ± SD: Slope | Std error of regression | $R^2$ | Regression coeff. ± SD: Intercept | Regression coeff. ± SD: Slope | Std error of regression |
| Tomato soups (A) | 11 | 0.68 | 4.42±0.09 | 0.059±0.014 | 0.21 | 0.81 | 4.41±0.07 | −0.0026±0.0004 | 0.16 |
| Tomato sauces, meatless (B) | 38 | 0.61 | 4.37±0.03 | 0.082±0.011 | 0.13 | 0.52 | 4.50±0.02 | −0.0019±0.0003 | 0.14 |
| Other tomato-vegetable comb. (C) | 26 | 0.82 | 4.36±0.03 | 0.116±0.011 | 0.12 | 0.77 | 4.52±0.03 | −0.0022±0.0002 | 0.14 |
| A + B | 49 | 0.67 | 4.39±0.03 | 0.067±0.007 | 0.15 | 0.68 | 4.50±0.02 | −0.0021±0.0002 | 0.15 |
| A + C | 37 | 0.66 | 4.43±0.04 | 0.072±0.009 | 0.18 | 0.78 | 4.50±0.03 | −0.0023±0.0002 | 0.14 |
| B + C | 64 | 0.71 | 4.36±0.02 | 0.100±0.008 | 0.13 | 0.68 | 4.51±0.02 | −0.0021±0.0002 | 0.14 |
| A + B + C | 75 | 0.67 | 4.40±0.02 | 0.076±0.006 | 0.16 | 0.72 | 4.50±0.02 | −0.0022±0.0002 | 0.14 |

[a] All values of correlation coefficient R significant at 0.01.

etable combinations, product categories representing many home canning recipes. As was true with the individual product categories, the acidity ratio and acidity difference for data set C were better predictors of product pH than was the titratable acidity, values of $R^2$ and the standard error for the latter index being only 0.34 and 0.24, respectively (data not included in Table 5). Consequently, titratable acidity was not considered further in this study as a predictor of the pH of combinations. pH predictions made for data sets A (tomato soups) and B (meatless tomato sauces) were less accurate than those made for data set C. Regression slopes differed substantially among these groups. The best pH predictions for soups were obtained with the acidity difference ($R^2 = 0.81$; std error = 0.16); the acidity ratio was a better pH predictor for meatless tomato sauces ($R^2 = 0.61$; std error = 0.13). The greater accuracy of pH predictions made with simple tomato-vegetable combinations than with soups and sauces is probably due to the more complex treatments received by the latter two groups of products during their preparation. Changes in composition affecting pH might occur as a consequence of the partial extraction of solubles from meat ingredients, bones, and vegetables that are strained out of the product following cooking. Heat-induced reactions and the concentration of organic acids, amino acids, and salts during prolonged cooking also might affect the product pH.

To further improve the accuracy of pH prediction, we again increased the number of data points by combining the pooled data sets (A, B, C) into new sets (A + B, A + C, B + C, A + B + C) from which new regression equations were obtained (Table 5). The substitution of the acidity difference equation for data set A + C in place of the equation derived from data set A to predict the pH of tomato-based soups resulted in the greatest improvement, reducing the standard error from 0.16 to 0.14. Combining B + C improved the accuracy of pH prediction for sauces. If a single equation for pH prediction were desired for all product categories, probably a marginal advantage with computerized computations, the acidity difference regression equation for data set A + B + C could be used.

Regression equations, derived from the pooled 1979 data sets, were validated by testing the goodness of fit of pH predictions made with product recipes taken from the 1981 data sets. Residuals between the predicted and experimentally determined pH values were not significantly different from residual mean square values obtained with the 1979 regressions, as judged by F-tests. Values of residual means for pH predictions made with 1979 and 1981 recipe data, using the 1979 regression equations, were no greater than ±0.02, indicating that the predictions are unbiased.

Having demonstrated the applicability of the 1979 regressions to the 1981 data, we recalculated the regression coefficients using the combined data sets. The new regression statistics were substantially the same as those given in Table 5, values of the standard error of regression being slightly smaller for the combined data sets. The new regression coefficients (Table 6) should be used in preference to those obtained from the smaller data set. Based on the correlation coefficients and standard errors obtained with these regressions, the best strategy for predicting the pH of tomato-based combinations would be as follows:

(a) For tomato soup recipes, use the acidity difference equation derived from data set A + C.

(b) For meatless tomato sauce recipes, use the acidity ratio or difference equation derived from data set B + C.

(c) For other tomato-vegetable combination recipes, use the acidity ratio or difference equation derived from data set C.

*—Continued on page 238*

*Table 6—Regression coefficients for pH prediction with the pooled 1979 and 1981 data sets*

| Pooled data set[a] | Recipes tested | Regression coefficient: Acidity ratio: Intercept | Acidity ratio: Slope | Acidity difference: Intercept | Acidity difference: Slope |
|---|---|---|---|---|---|
| A | 11 | 4.42 | 0.0591 | 4.41 | −0.00262 |
| B | 56 | 4.38 | 0.0813 | 4.50 | −0.00174 |
| C | 38 | 4.35 | 0.1185 | 4.53 | −0.00211 |
| A + B | 67 | 4.40 | 0.0664 | 4.50 | −0.00199 |
| A + C | 49 | 4.40 | 0.0764 | 4.51 | −0.00214 |
| B + C | 94 | 4.36 | 0.1006 | 4.52 | −0.00199 |
| A + B + C | 105 | 4.39 | 0.0780 | 4.51 | −0.00204 |

[a] As defined in Table 5.

(d) If a single equation were required for all tomato-based combination recipes, use the acidity difference equation derived from data set A + B + C.

The goodness of fit of data points to the last regression equation (acidity difference, data set A + B + C) is shown in Fig. 1. The confidence limits represent two standard errors. The standard errors for the regression equations recommended above (0.12–0.13 pH unit) are slightly larger than those associated with the model (0.11 pH unit) reported previously (Sapers et al., 1982b). Although the model may be a marginally better predictor of product pH, it would not be applicable to about 40% of the recipes in our compilation because of its restriction to specific ingredients. In contrast, the new equations have no such limitation and could be applied to virtually all recipes for tomato-based combinations.

In making use of the new equations to determine whether a recipe yields a high-acid or low-acid product, certain precautions stated previously in connection with the use of the model apply here as well. First, regression equations developed for specific product categories should not be applied to other product categories outside the scope of this study. Second, because of the inherent uncertainty of statistical predictions, confidence limits agreed upon by USDA Extension in consultation with food preservation specialists should be placed around predicted pH values. If the confidence limits for a given recipe included pH 4.6, the product made by that recipe should be considered a potential low-acid food. Third, recipes called into question because of high predicted pH values should be tested with typical ingredients, avoiding the use of tomatoes or other ingredients that may be atypical with respect to variety, maturity, ripeness, or condition, causing the combination to be unusually high or low in acidity.

## REFERENCES


DiVito, A.M., Sapers, G.M., and Phillips, J.G. 1982. Acidification requirements for home canned combinations of tomatoes and low-acid ingredients. J. Food Sci. 47: 2089.

Harrel, C.G. and Thelen, R.J. 1959. "Conversion Factors and Technical Data for the Food Industry," 6th ed., p. 474. Burgess Publishing Co., Minneapolis, MN.

Institute of Food Technologists (IFT). 1977. Home canning, a scientific status summary. Food Technol. 31(6): 43.

Ito, K.A. and Chen, J.K. 1978. Effect of pH on growth of *Clostridium botulinum* in foods. Food Technol. 32(6): 71.

Odlaug, T.E. and Pflug, I.J. 1978. *Clostridium botulinum* and acid foods. J. Food Protection 41: 566.

Paulson, K.N. and Stevens, M.A. 1974. Relationships among titratable acidity, pH and buffer composition of tomato fruits. J. Food Sci. 39: 354.

Sapers, G.M., Phillips, J.G., and DiVito, A.M. 1982a. Equilibrium pH of home canned foods comprising combinations of low-acid and high-acid ingredients. J. Food Sci. 47: 277.

Sapers, G.M., Phillips, J.G., DiVito, A.M., and Brooks, W.M. 1982b. Model for predicting the pH of foods comprising mixtures of tomatoes and low-acid ingredients. J. Food Protection 45: 566.

Sapers, G.M., Phillips, J.G., and Stoner, A.K. 1977. Tomato acidity and the safety of home-canned tomatoes. HortScience 12: 204.



Ms received 6/24/83; revised 8/26/83; accepted 9/12/83.

The authors thank Sandra P. Graham, Elinor Gimbel, and Lucille K. Conway, employees of ERRC, for their technical assistance.