CLASSIFICATION OF THE CHARACTERISTICS OF TWO MANGO CULTIVARS
HARVESTED AT DIFFERENT STAGES OF MATURITY USING GAS
CHROMATOGRAPHY AND SENSORY DATA

By

Yael Vodovotz

B. Sc. (Agr.) University of Illinois, 1989

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in
THE FACULTY OF GRADUATE STUDIES
Department of Food Science

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 1992

© Yael Vodovotz, 1992
In presenting this thesis in partial fulfilment of the requirements for an advanced
degree at the University of British Columbia, I agree that the Library shall make it
freely available for reference and study. I further agree that permission for extensive
copying of this thesis for scholarly purposes may be granted by the head of my
department or by his or her representatives. It is understood that copying or
publication of this thesis for financial gain shall not be allowed without my written
permission.

(Signature)

Department of Food Science
The University of British Columbia
Vancouver, Canada

Date April 8, 1992
ABSTRACT

The classification of the characteristics of two Mexican mango varieties (Tommy Atkins and Haden) ripened under two conditions (in cold storage and on the tree) was attempted using several multivariate analysis techniques. The aroma of the different mango purees was analyzed using both capillary gas chromatography and sensory analysis. A new portable gas chromatograph (SRI model 8610, Torrance, CA) equipped with a purge and trap was used and the area count of relevant peaks calculated. These results and those obtained from a sensory panel comprised the classification factors. In addition, parameters such as the Magness Taylor pressure values, pH and acid/sugar ratio were determined. Two familiar classification techniques namely Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) were used as well as a new program developed in our laboratory: PCA-Similarity (PCA-SIM). Results showed that differences exist between varieties as well as ripening conditions when analyzed both by GC and sensory techniques. Distinct groups were formed when the results of these techniques were independently subjected to LDA and PCA-SIM. PCA was more successful as a variable reduction technique than a classification method. Although both GC and sensory methods were successful in characterizing the differences between the groups, one could not replace the other since no correlation was found between the two methods. The other physico/chemical parameters were useful, but only to
a limited extent, and none could account for both differences between varieties and ripening conditions.
TABLE OF CONTENTS

ABSTRACT........................................................................................................... ii
TABLE OF CONTENTS............................................................................................. iv
LIST OF TABLES...................................................................................................... vii
LIST OF FIGURES..................................................................................................... viii
LIST OF APPENDICES............................................................................................ xii
ACKNOWLEDGEMENT............................................................................................. xiii
I. INTRODUCTION................................................................................................. 1
II. LITERATURE REVIEW.......................................................................................... 4
   A. Maturity of Mangos......................................................................................... 4
   B. Pretreatment of Mangos................................................................................ 6
   C. Storage of Mangos........................................................................................ 8
   D. Flavour of Mangos....................................................................................... 9
   E. Gas Chromatography..................................................................................... 12
      1. The Column.................................................................................................. 12
      2. Traps............................................................................................................. 14
      3. Detectors..................................................................................................... 15
      4. Sampling Procedures.................................................................................. 16
      5. Temperature Programming........................................................................ 18
   F. Optimization.................................................................................................... 19
      1. Random Centroid Optimization............................................................... 20
   G. Sensory Evaluation...................................................................................... 21
III. MATERIALS AND METHODS........................................................................... 22
   A. Samples.......................................................................................................... 22
      1. For Optimization of GC Parameters......................................................... 23
      2. For Actual GC Experimentation............................................................... 23
3. For Sensory Analysis..........................24
4. For GC/MS.....................................24

B. Sample Preparation for Analysis...............25
   1. Samples Ripened in Storage...............25
   2. Gas Chromatography.......................27
   3. Samples Ripened on the Tree and Those
      Used for Optimization of the GC
      Parameters................................30
   4. GC/MS....................................30
   5. Titratable Acidity and pH...............31
   6. Sugar Analysis............................31
   7. Sensory Analysis..........................32
   8. Magness Taylor Pressure Determination....35

IV. RESULTS AND DISCUSSION..........................37
   A. Gas Chromatography.........................37
      1. Optimization: Preliminary..............37
      2. Optimization: Random Centroid
         Optimization..........................52
      3. Internal Standard........................60
      4. Repeatability of Chromatograms.........61
      5. Comparison of Peaks and their Selection
         for Multivariate Analysis...............61
      6. Effect of Fruit Size....................64
   B. GC/MS....................................67
   C. Multivariate Analysis Techniques..........73
      1. Principal Component Analysis............73
      2. Linear Discriminant Analysis............74
      3. Principal Component Analysis-
D. Multivariate Analysis of GC Results..........................79
1. Principal Component Analysis.........................79
2. Linear Discriminant Analysis.....................82
3. Principal Component Analysis-Similarity....................84
E. Sensory Analysis Results.................................89
1. Taste.............................................89
2. Aroma.............................................89
F. Multivariate Analysis of Sensory Analysis Results...............103
1. Principal Component Analysis..................103
2. Linear Discriminant Analysis................105
3. Principal Component Analysis-Similarity..................108
G. Comparison of Chemical and Sensory Data for Taste..........................111
H. Comparison of Sensory and GC Data......................113
I. PH and Magness Taylor Results......................113
J. Weight Loss in Mangos..................................119
V. CONCLUSION........................................121
VI. REFERENCES.....................................126
VII. APENDIX........................................134
LIST OF TABLES

Table                                      Page

1. Experimental plan generated by Random Centroid Optimization (RCO) ........................................55

2. The mean, standard deviation and coefficient of variation of 5 peaks obtained from each of 3 Tommy Atkins chromatograms run under optimized conditions to test the repeatability of the GC .......................................................... 63

3. Results of PCA of the seven normalized peak areas for the 112 gas chromatograms .......................... 80

4. Actual group membership (rows) vs. predicted (columns) for LDA of GC results using both individual variables and the first 3 principal components from PCA .................................................. 83

5. Averaged and normalized sensory results for 16 mango samples ...................................................... 90

6. Averages and standard deviation of the normalized results for the nine sensory attributes .............. 104

7. Results of PCA of the sensory attributes .......... 106

8. Actual group membership (rows) vs. predicted (columns) for LDA of the individual descriptive terms as well as principal components from PCA .......... 109

9. Averages for each lot of pH and Magness Taylor values ................................................................. 118

10. Average weights and standard deviations of mangos before and after storage (for those ripened in storage) and after picking (for those ripened on the tree). The % weight loss was calculated for mangos ripened in storage ................................................. 120
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The SRI model 8610 gas chromatograph</td>
<td>28</td>
</tr>
<tr>
<td>2. Questionnaire for aroma determination of mango samples using QDA</td>
<td>34</td>
</tr>
<tr>
<td>3. Questionnaire for sweet/sour determination of mango using QDA</td>
<td>36</td>
</tr>
<tr>
<td>4. Blank run (10 mL distilled water) of a Tenax trap after 15 min of baking at 230°C</td>
<td>39</td>
</tr>
<tr>
<td>5. Blank run (10 mL distilled water) of a Tenax trap after 30 min of baking at 230°C</td>
<td>41</td>
</tr>
<tr>
<td>6. Five gram sample of Tommy Atkins heated to 45°C for 15 min. (dry purge 19 min.), concentrated on a Tenax trap and eluted onto a capillary DB-624 column</td>
<td>42</td>
</tr>
<tr>
<td>7. Tommy Atkins sample (2.5 grams) with the same conditions as in Figure 6</td>
<td>43</td>
</tr>
<tr>
<td>8. Tommy Atkins sample (0.50 grams) with the same conditions as in Figure 6</td>
<td>44</td>
</tr>
<tr>
<td>9. Tommy Atkins sample (0.10 grams) with the same conditions as in Figure 6</td>
<td>45</td>
</tr>
<tr>
<td>10. Tommy Atkins sample (0.25 grams) heated to 40°C for 15 min., concentrated on a Tenax trap and eluted onto a capillary DB-WAX column</td>
<td>46</td>
</tr>
<tr>
<td>11. Tommy Atkins sample (0.10 grams) purged at 28°C for 15 min., concentrated on a Tenax trap and desorbed onto a capillary DB-WAX column with a Helium flow rate of 1.96 mL/min</td>
<td>48</td>
</tr>
<tr>
<td>12. Tommy Atkins sample with the same conditions as in Figure 11 except volatiles were concentrated on a charcoal trap</td>
<td>49</td>
</tr>
<tr>
<td>13. Tommy Atkins sample (0.50 grams) purged at 39°C for 12 min., concentrated on a charcoal trap and desorbed onto a Megabore DB-624 column with a Helium flow rate of 4.62 mL/min. The temperature program was modified to ramp at 2 deg/min instead of 4 deg/min</td>
<td>50</td>
</tr>
</tbody>
</table>
14. Tommy Atkins sample with the same conditions as in Figure 13 except sample was desorbed onto a Tenax trap.................................51

15. The resolution between two peaks was calculated by dividing the value of $f$ by $g$. The resulting $P_i$ value contributed to the Chromatographic Response Function.......................................54

16. Mapping results of the first cycle of experiments generated by RCO..................................................56

17. Mapping results of all 18 experiments generated by RCO.................................................................58

18. Tommy Atkins mango chromatograms obtained using the RCO program.............................................59

19. Reproducibility of three chromatograms of Tommy Atkins using the optimized conditions.............62

20. Sample chromatogram of Tommy Atkins ripened on the tree.........................................................65

21. Sample chromatogram of Haden ripened in storage.....66

22. Simulated chromatograms of lot 1 (size 12)= A and lot 3 (size 16)= B, of Haden mangos utilizing adjusted areas of the seven peaks used for multivariate analysis.................................68

23. Simulated chromatograms of lot 2 (size 14)= A and lot 4 (mixed sizes)= B, of Tommy Atkins mangos utilizing adjusted areas of the seven peaks used for multivariate analysis.............69

24. Mass spectrometry identification of major peaks of Tommy Atkins sample.................................70

25. Mass spectrometry identification of major peaks of Haden sample..................................................71

26. Mass spectrometer spectra and chemical structure of $\alpha$-Pinene, Car-3-ene and $\beta$-Phellandrene......72

27. Flow diagram of Principal Component Analysis-Similarity program....................................................77
Figure | Page
--- | ---
28. Plot of PCA factor scores (principal components), derived from GC results for four representative lots. | 81
29. Plot of factor scores (canonical variables) resulting from LDA for the individual peak areas in the GC analysis. | 85
30. Principal Component Analysis - Similarity results of the seven GC peaks utilizing Tommy Atkins, lot 1, individual 4, ripened in storage as a standard. | 87
31. Corrected principal component scores of the standard (Tommy Atkins, lot 1, individual 4, ripened in storage) compared to the corrected principal component scores of Haden, lot 2, individual 3, ripened in storage. | 88
32. Averaged and normalized panel ratings for sweetness and sourness using QDA for Tommy Atkins and Haden. | 91
33. Box plots of sensory data before and after normalization procedure for the parameter fruity. | 93
34. Box plots of sensory data before and after normalization procedure for the parameter weed. | 94
35. Box plots of sensory data before and after normalization procedure for the parameter sweet aroma. | 95
36. Box plots of sensory data before and after normalization procedure for the parameter pineapple/banana. | 96
37. Box plots of sensory data before and after normalization procedure for the parameter honeydew melon. | 97
38. Box plots of sensory data before and after normalization procedure for the parameter overall intensity. | 98
39. Box plots of sensory data before and after normalization procedure for the parameter sour taste. | 99
<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>40. Box plots of sensory data before and after normalization procedure for the parameter tea</td>
<td>100</td>
</tr>
<tr>
<td>41. Box plots of sensory data before and after normalization procedure for the parameter sweet taste</td>
<td>101</td>
</tr>
<tr>
<td>42. Explanation of the Box plot structure and symbols</td>
<td>102</td>
</tr>
<tr>
<td>43. Sensory PCA results of principal component 1</td>
<td>107</td>
</tr>
<tr>
<td>44. Plot of factor scores (canonical variables) resulting from LDA for the individual sensory descriptive terms</td>
<td>110</td>
</tr>
<tr>
<td>45. Scattergrams of individual peak areas (PA1-PA7) from GC analysis with individual sensory descriptive terms: an attempt to find a correlation between sensory and GC</td>
<td>114</td>
</tr>
<tr>
<td>46. Scattergrams of factor scores from PCA of GC results with individual sensory descriptive terms: an attempt to find a correlation between sensory and GC</td>
<td>115</td>
</tr>
<tr>
<td>47. Magness Taylor and pH results of the samples used for sensory analysis</td>
<td>116</td>
</tr>
</tbody>
</table>
VII. LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Principal Component Analysis - Similarity computer program</td>
<td>134</td>
</tr>
<tr>
<td>2. Explanation of symbols for each mango sample</td>
<td>141</td>
</tr>
<tr>
<td>3. Sample calculation for the box plot</td>
<td>142</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENT

I would like to thank Dr. S. Nakai for sponsoring this project: his vision and steadfast support were vital in its completion. I am grateful to Belinda Cordoba and Guillermo Arteaga for everything from scientific to moral support; without them this project would have never got off the ground. In Mexico, I would like to thank Dr. H. Weinstein and his wife for their gracious hospitality and Ing. Balleza and his family for all their kindness and unquestionable trust. To the Fernandez brothers for the cooperation of their employees, the use of their processing plant and continual help, thank you. Last, but far from least, my gratitude to my parents, family and many friends for being there when I needed them most.
I. **INTRODUCTION**

Mangos, although not grown in Canada, are becoming widely accepted by the public in the "exotic" fruit category. Aside from vitamins A and C, mangos provide minor amounts of nutrients and are, therefore, considered a secondary food source providing variety and flavour to the diet (Caygill et al., 1976). This fruit can, in fact, be found in most food markets in the Vancouver area. Although mangos are grown in many tropical and subtropical regions of the world (i.e. India, Australia, Africa, South and Central America as well as parts of North America), the main suppliers to the Vancouver market are located in Mexico and the Philippines. Aside from different growing regions, there are hundreds of varieties available and certainly more than one grown in each region.

The two most popular cultivars supplied from Mexico are Haden (available from May to mid July) and Tommy Atkins (available in July and August) which are grown in the Pacific Coast region. The Tommy Atkins is a medium to large size fruit weighing 450-700 grams. This cultivar is oval in shape with a broadly rounded tip. The skin is smooth and thick which protects it from disease. The external color is orange-yellow with a bright red blush. The yellow flesh is substantially fibrous which may be objectionable to some (Campell, 1973). The Haden variety is similar in size, but is much less fibrous and lacks the attractive colouring of the Tommy Atkins.
Because of the long transport time required for fruit from Mexican fields to reach Canadian markets, mangos need to be harvested before reaching the final ripening stages. The fruit is then ripened during shipment and, ideally, these mangos that are ripened in storage should be similar in composition to those ripened on the tree.

Medlicott and co-workers in 1988, discussed the difficulties in obtaining harvested mangos that will fully ripen simultaneously during transportation and the problems of different degrees of maturity in retail markets. Resultant degree of maturity is an important parameter since consumers expect the pleasing aroma and savory taste associated with properly ripened fruit. There is, however, no accurate method for determining harvest maturity. Some physical parameters that have been studied for this purpose are: size, shape, peel and flesh colour, lenticels, shoulder growth, pit around the pedicel, specific gravity and heat units. Chemical composition parameters such as starch content, sugar/acid ratio and phenolic content may be of value in assessing mango maturity. However, none of the above factors have been found to be reliable for this purpose (Salunkhe and Desai, 1983). Suryaprakasa Rao in 1972 also found no relationship between degree of ripeness and the size of the fruit and concluded that the principal physical components of the fruit were independent of its maturity.

Since, in the case of fruits, the aroma quality
constitutes the main portion of the overall sensory quality (Schamp et al., 1982), mango volatiles were analyzed. Gas chromatography (GC) is a common method used to detect volatile compounds in food. As mentioned above, consumer acceptance is due, partly, to the proper aroma development which may be indicative of maturity. Gas chromatography, however, has not been widely used for this purpose in mangos.

The objectives of this research were the following:

1. To optimize the conditions of a new, portable, GC for the headspace analysis of mango puree.
2. To group mangos according to cultivar and ripening conditions using results obtained from gas chromatography and sensory analysis.
3. To compare a computer program developed in this laboratory (PCA-Similarity) to other existing multivariate analysis techniques in the classification of mango quality.
4. To relate physical and chemical properties of mangos to cultivar quality and ripening conditions.
II. LITERATURE REVIEW

A. MATURITY OF MANGOS:

The maturity stage of harvested mangos has a tremendous influence on the quality of the storage-ripened fruit. Mature but unripe fruit has high total acidity and ascorbic acid values but low contents of sugars and total soluble solids (Askar et al., 1983). Mangos that were harvested at the semi-ripe stage had a longer storage life than mangos harvested fully ripe but did not ripen evenly nor had the same composition than fruit that was harvested when fully mature (Medlicott et al., 1990). Also, fruit harvested too early was subject to heavy spoilage during storage (Lakshminarayana et al., 1974). Therefore, mangos, regardless of cultivar or origin, should not be harvested before they reach full maturity. Harvesting of semi-ripe mangos usually occurs when the skin colour just begins to change and one or two fruits drop from the tree (Jain, 1961).

A great deal of research has been carried out on determining and controlling the ripening process of mangos. Mangos are a climacteric fruit and Krishnamurth et al. in 1970 found that colour break, odour development and softening coincided with the climacteric peak of advanced maturity in Pairi mangos. Lakshminarayana and co-workers in 1970 noted that the growth climacteric may be due to high enzymatic activity in the initial stages when the fruit is utilizing
enormous quantities of respiratory substrates.

In Haden mangos, the rate at which flesh firmness was lost declined after the first three days as measured by a Magness Taylor Pressure tester with a plunger of 5/16 inches diameter. Likewise, a sensory panel found that the mangos became acceptable after the third day of storage (Soule et al., 1956). Density was found to be directly correlated with firmness for various Indian mango cultivars (Verma and Bajapai, 1971). Yet none of the aforementioned parameters have been successfully quantified so as to allow the shipper to use them as selection tools for the desired maturity.

Bandyopadhyay and Gholap in 1973 examined the relationship of quantity and type of fatty acid and degree of maturity of mangos. Their experiments with Alphonso mangos revealed that the degree of unsaturation of the fatty acids became greater as fruits matured and that these UFA were more metabolically active during ripening. Also, the ratio of palmitic to palmitoleic acid was found to be significant in contributing to the aroma and flavour of the mango. Strong aroma was noted when the ratio was smaller than one and a mild aroma when the ratio was greater than one.

Other parameters which were investigated for their relationship to maturity include starch content and amylase activity. Pandey and co-workers in 1974, found that the greatest change in carbohydrates during the development of Dashehary mangos was an increase of starch content from one to
13%, a result also found by Pompanoe and co-workers (1957). They concluded that the analysis of the flesh for starch content would be useful to estimate the degree of maturity of the hard green fruit. Amylase activity was also found to increase with the development of the fruit (Fuchs et al., 1980) and this increase was parallel to the increase of fruit weight (Fuchs et al., 1980 and Tandon and Kalra, 1983).

Little research has been carried out on comparing tree-ripe mangos to those harvested mature but green and then ripened in storage. Askar and co-workers (1983) compared these two conditions for the Egyptian mango variety, Zebda. They concluded that tree ripened fruits were better in taste and colour than those ripened in storage at room temperature. Also, tree ripened fruits had more aroma compounds as compared to those ripened in storage especially in the concentration of cis-octimene, a very important compound in Egyptian mango aroma.

B. PRETREATMENTS OF MANGOS:

A common problem encountered when trying to extend the storage life of tropical fruit is chilling injury. This phenomenon occurs when a fruit is stored at temperatures below its tolerance level (Mukerjee and Srivastava, 1979). Chilling injury, according to Thomas and Oke (1983) is manifested by pitting and browning of skin tissues, uneven ripening upon the removal from low temperatures and failure to develop normal
aroma, flavour, skin and flesh colour on ripening and decreased resistance to fungal diseases. Chilling injury may be caused by a low production of certain essentials or an over production of some toxic products (Salunkhe et al., 1983). Farooqi and co-workers (1985), while investigating chilling injury in Sensation and Samor Bahisht mango cultivars found that the skin showed the symptoms of chilling injury while the rest of the fruit remained unaffected. Mature or partially ripe Haden mangos developed chilling injury whereas ripe mangos did not at the same temperature (Mukerjee and Srivastava, 1979).

Several pretreatments had been suggested to prevent spoilage, promote uniform ripening and avoid chilling injury in mangos. Precooling, both by dipping the mature green mangos in cold water for one minute (Mann and Singh, 1975) and cold adaptation by slow cooling from 10.5 C to 5.5 C (Mukerjee and Srivastava, 1979) resulted in less spoilage of the fruit. However, the most successful treatment has been a hot water dip (50 C for 7 min) prior to storage to decrease fungal decay (Thomas, 1975), especially in Haden mangos which are very sensitive to anthracnose spoilage. The hot water treatment removed part of the natural waxes on the fruit surface and facilitated the exchange of respiratory gases by the increased cell wall permeability. However, this treatment resulted in reduced storage life due to earlier ripening and slightly greater weight loss of the fruit (Salunkhe and Desai, 1983).
C. STORAGE OF MANGOS:

The optimum conditions of storage for mangos are highly dependant on the particular cultivar as well as the degree of maturity of the fruit. Storage temperature was considered the most influential part on the ripening time (Fornaris-Rollan et al., 1990), and various authors have looked at the influence of storage temperature on mangos (Tripathi, 1988; Krishnamurthy and Joshi, 1989; Medlicott et al., 1990; Thomas, 1975; Kapse et al., 1985; Abou Aziz et al., 1975 and Saucedo Veloz et al., 1977). Lowering storage temperature slowed down the physiological and biochemical activities of the fruit leading to senescence (Mukerjee and Srivastava, 1979). Higher acidity and a slower increase in sugar content resulted when Pairi mangos were held at 0 C for various number of days. Campbell in 1959 summarized the findings by recommending storage temperatures of 18-24 C at 85-90% relative humidity. Below 10 C chilling injury was common and above 27 C quality was seriously impaired. Similarly, Salunkhe and co-author in 1983, recommended thee storege of mangos above 19 C since lower temperatures caused a greyish scald-like discoloration of the skin and uneven ripening, a symptom of chilling injury.

Holding the mangos at temperatures above 36 C resulted in over-ripe fruit which rotted quickly (Patwardhan, 1927). Bandyopadhyay et al., in 1973 found that ripening of Alphonso mangos at 25-30 C was accompanied by a change in fatty acid composition as well as an increase in triglyceride content in
the pulp, while ripening at low temperatures did not result in such changes. Saha et al., in 1976 demonstrated the effect of ripeness level on both physical appearance and chemical composition of the mangos.

D. FLAVOUR OF MANGOS:

Despite its large production volume and increasing economic importance, mango and its flavour characteristics have not been widely studied (Shibamoto and Tang, 1990).

Researchers who had analyzed flavour from the perspective of taste had utilized the sugar/acid ratio in their work. The predominant acid found in mangos was citric with lesser amounts of succinic and malic acids (Medlicott and Thompson, 1985). The titratable acidity was shown to decrease as ripening progressed (Saha et al., 1976; Selvaraj and Pal, 1985; Sharaf et al., 1989; Askar et al., 1983; Pandey et al., 1974; Suryaprakasa Rao et al., 1972; and Mukerjee, 1959.) Tripathi in 1988 found that for mangos stored at room temperature, the acidity decreased for the first six days but gradually increased thereafter.

The three predominant sugars in mango are sucrose, fructose and glucose. Their abundance was again highly dependant on the maturity of the fruit. Furan derivatives produced by the Maillard reaction were found in abundance in mangos and probably contributed to its sweetness (Sharaf et al., 1989.) Lakshiminarayana and co-workers in 1970 observed
that total sugars increased gradually up to harvest maturity, but a slight decline was observed near ripening. A similar trend was observed by Patwardhal in 1927 and by Mukerjee in 1959. Medlicott and Thompson in 1985 agreed with Fuchs and co-workers who in 1980 explained this phenomenon was caused by the depletion of starch which initially replaces the sugar being metabolised.

Upon maturation, the decrease in acidity and increase in sugars resulted in a higher sugar/acid ratio and this balance determined, to a great extent, the taste of the mango (Selvaraj and Pal, 1989; Sharaf et al., 1989) and cultivars with higher ratios were preferred by most consumers (Lodh et al., 1974). Suryaprakasa Rao and co-workers in 1972, however, could find no relationship between the maturity of mango and the sugar/acid ratio.

Some research has been conducted on identifying the compounds which contribute to mango aroma, mostly utilizing GC/MS techniques. Unlike most other fruits, each cultivar of mango differs greatly in its chemical composition as well as physical appearance. Each cultivar has its own unique flavour and aroma, and even though, monotropene hydrocarbons are the major group of volatiles in mango aroma, there is no single chemical known to have a characteristic mango odour (Shibamoto and Tang, 1990). Engel and Tressl in 1983 found that limonene was the main component in Baladi mango but was present in very minute quantities in the Alphonso variety. Also, MacLeod and
Pieris in 1984 found that Venezuelan mangos contained mainly car-3-ene. Abd El-Baki and co-workers in 1981, found that cis-ocimene was a character impact compound in mango aroma.

Difference in the aroma profile are dependent on the maturity of the mango. Ackerman and Traline in 1984 found that 2-butanoic acid esters were characteristic of green mango aroma while 3-butanoic acid esters comprised ripe mango aroma. These facts contribute to the difficulty of identifying and characterizing typical mango flavour. Additionally, most methods of isolation of the volatiles consisted of distillation and organic solvent extraction (Engel and Tressl, 1983; Yamaguchi et al., 1983; MacLeod and Pieris, 1984; Hunter et al., 1974; Gholap and Bandyopadhyay, 1977; Idstein and Schreier, 1985; Gholap et al., 1986; Kunishi and Seale, 1961; Abd El-Baki et al., 1981; Pino et al., 1989, Franco et al., 1991 and Sakho et al., 1985). These authors studied many cultivars (Alphonso, Baladi, Parrot, Willard, Jaffna, Keitt, Tommy Atkins, Zebda, Fairly, Corazon, Bizcochuelo, Super Haden, Haden Edward, Palmer and Zill) grown in several countries (Egypt, Brazil, Philippines, Venezuela, Sri Lanka, India and Mexico). The major components identified (and here again largely dependent on the cultivar) belonged to the following groups: hydrocarbons, esters, alcohols, aldehydes, ketones, lactones and terpenes.

Some criticism has surfaced as to the use of these methods for volatile compound isolation. Bartley and Schwede
in 1987, warned that both qualitative and quantitative changes may occur to the volatile components when solvent extraction and distillation were used. In addition, a loss of volatile components may take place due to the stripping away of low boiling point compounds (Ackerman and Troline, 1984). Headspace concentration was recommended as more meaningful method than distillation or extraction to ascertain volatiles released from fruits (Schamp and Dirinck, 1982).

E. GAS CHROMATOGRAPHY:

Before the analysis of the mango samples could be carried out, there were various parameters of the GC which needed to be considered in order to obtain the most information in a reasonable time span.

1. The Column

The column is probably the most important part of a GC since that is where separation of the volatiles occurs (Shibamoto, 1984). Many different columns are available so proper choice is based on several parameters. The first is the type of the column, which is dictated by the type and thickness of the stationary phase, column diameter and column length. The type of stationary phase depends on its bonded phase which ranges from nonpolar to polar. Nonpolar columns separate on the basis of boiling point alone and, therefore, are rarely used for complex mixtures such as food. Polar columns separate not only on the basis of boiling point but
also on the basis of chemical functionality; however, they are extremely sensitive to oxidation and have a low tolerance for abuse. A medium polarity column resolves a large number of compounds with similar structure and combines the benefits of both polar and nonpolar phases (Rood, 1991).

The diameter of a column is the main factor which determines capacity: the maximum amount of sample that can be injected onto a column before significant peak distortion occurs (J&W, 1990). The greater the diameter the larger the capacity. However, capacity is often sacrificed in order to obtain maximum resolution as is the case with capillary columns (0.35mm ID) as compared with packed columns (16mm ID). A good compromise is the megabore columns which have a slightly larger internal diameter than capillary columns (0.53mm) and have a substantially larger capacity. The thickness of the stationary phase also affects the capacity and resolution (both these parameters would increase with a thicker film). Greater film thickness increases the column’s inertness and retention and enables higher elution temperatures which may eliminate the need for expensive subambient operation. However, very thick films also result in large amounts of column bleed.

The length of the column affects the resolution: the longer the column the greater the resolution. However, to double the resolution between two solutes, the column length must increase by a factor of 4 which results in very long

2. Traps

For the analysis of trace compounds in a headspace, a concentration step usually precedes the actual analysis. Several methods are used such as cryogenic concentration and trapping on adsorbent materials. Using cryogenics results in excellent volatile recovery but the cost of operation is sometimes prohibitive especially if liquid nitrogen is used. However, cryo-trapping during purging may be a cheap alternative since liquid CO$_2$ can be used. Cryo-trapping differs from cryo-focusing in that only the trap is cooled and not the GC oven. The adsorbent traps which are often used are coated with porous polymers which have little affinity for water but high affinity for organic compounds (Durr, 1983). Recovery of the volatiles on adsorbents may be achieved by elution with solvents or by thermal desorption. Activated charcoal, and Tenax (p-2,6-diphenyleneoxide) are commonly used traps. The activated charcoal trap has very strong adsorbing capacity and can be heated to high temperatures but its desorption characteristics are very poor. Tenax is by far the most common adsorbent in use because of its excellent trapping of non-polar compounds, its inability to retain water as well as efficient desorption with very low bleed levels. However, polar compounds are poorly retained and because of its low
desorption temperature limit, heavier compounds may remain in

3. Detectors

There are several types of detectors that may be used
with the GC. Sensitivity- the minimum amount of a compound
detected, selectivity- the type of compound detected and
linear range- the detector response proportional to compound
concentration are important parameters that need to be
considered when choosing an appropriate detector for a
particular application (Rood, 1991).

The thermal conductivity detector (TCD) is sensitive to
a change in resistance caused by cooling of a heated wire.
The quantity of heat removed from the wire depends on the
velocity and mass of a molecule. This detector is sensitive
to all organic compounds and accommodates a variety of carrier
gasses including helium, hydrogen and nitrogen (Pomeranz and
Meloan, 1978). However, because it responds to water and
other non-hydrocarbons, it is usually used only when other
more sensitive detectors are unsuccessful (SRI, 1991). Also,
its lack of sensitivity to certain compounds makes it
inadequate for many applications (Durr, 1883).

Another type of detector is the flame ionization detector
(FID) which operates on the basis of a flame produced from a
combination of hydrogen, air and the carrier gas. The flame
(2100 C) burns the sample producing negative and positive ions
which create a current leading to a detectable signal. The FID responds to any compound with a carbon-hydrogen bond and is 1000 times more sensitive than a TCD leading to its more frequent use (Pomarenz and Meloan, 1978). This detector is also very stable since its calibration stays constant as long as the gas flow rates don’t change (SRI, 1991).

An electron capture detector (ECD) may be used for added selectivity. In this detector the electrons striking the sample molecules are captured decreasing the original signal. It can, therefore be used for halogenated compounds, conjugated carbonyls, nitriles, nitrates and some organometallic compounds (Pomeranz and Meloan, 1978). However, the ECD is radioactive and is prone to contamination by dirty samples and air leaks (SRI, 1991).

4. Sampling Procedures

There are two main types of sampling techniques available for GC analysis. The first is that of injection of the actual sample which is vaporized in the injector. Although this method reveals the content of the sample it is not always an accurate representation of what the nose smells. Difficulties also arise when the sample is solid. There are four main injection techniques: direct, split, splitless and cool on column. In direct injection, the entire sample is completely transferred onto the column while in split and splitless injection only a fraction of the sample reaches the column.
Split injection is used for very concentrated samples since only a small fraction is allowed to enter the column while in splitless most of the sample reaches the column. The splitless mode is often preferred to the split injection due to lower discrimination against high boiling, less volatile compounds. The cool on column technique involves the deposition of the sample directly on the column without vaporization and is, therefore, used for trace level, high boiling or thermally unstable samples (Rood, 1991).

If extraction with organic solvents is used, artifacts may be formed as well as low boiling point compounds may be lost (Dirinck et al., 1984). Quantitative and qualitative changes in the volatile composition occur during distillation and extraction and these methods do not take into account the relative volatility of the flavour compounds (Schamp and Dirinck, 1982). For example, a much larger number of peaks were eluted from a commercial room freshener extracted with pentane than from the head space sampling of the same sample (Shibamoto, 1982). Head space sampling, therefore, is considered a much better method since it measures the volatiles released from the food matrix which can be better correlated to sensory analysis. Headspace analysis does not destroy the sample which may be of some advantage in certain applications (Dirinck et al., 1984). It is considered an ideal method for sample examination (Durr, 1983).

The method of sample delivery and acquisition is also
important. If syringes fitted with steel needles are used they act as a miniature fractionation system and do not always deliver the full sample causing poor reproducibility and inaccurate results (Jennings and Takeota, 1984). Two alternatives are available: static (equilibrium) headspace analysis and dynamic headspace analysis. In the static method the volatiles that are in equilibrium with a sample in a closed system are measured. This system has the benefit that it can be easily automated, however, only the more volatile compounds are detected and large sample volumes are necessary for proper analysis. In the dynamic headspace technique, a continuous flow of gas is passed through the sample to remove the volatiles which are then concentrated on an adsorbent polymer trap (Dirinck et al, 1984). The purge and trap method is very sensitive and yields reproducible results and is used for samples comprised of a large number of compounds with closely related structures. A major problem with this technique is the large quantity of water which is generated and introduced into the column extinguishing the FID detector. Reducing the purge time and desorption time may solve this problem.

5. Temperature Programming

Temperature programming (changing the oven temperature with time) is considered the most influential parameter on the selectivity and proper separation of an analysis (Schomburg et al, 1984). If isothermal temperature is used, low boiling
point compounds will elute with higher boiling point species resulting in broad bands with little to no separation. The exact temperature program used would highly depend on the volatiles present.

F. OPTIMIZATION

After these parameters are established, an optimization procedure should follow to obtain the best results given the limitation of the GC when using purge and trap for sampling. Optimizing the purge conditions such as purge volume, sample size and sample temperature increases the sensitivity of the system and is, therefore, widely used. Although intuitively all three factors should be increased, doing so would lead to poor results. For example, sample size is often increased, but response is not linear with sample size unless the purge volume is adjusted, especially for compounds with a low purge efficiency. Also recovery of some volatiles may actually decrease due to breakthrough on the trap (caused by supersaturation of the trap). Increasing the temperature of a sample will increase recovery of the volatiles due to higher vapour pressure and in solids greater mobility in the matrix. Higher temperature will, however, lead to a greater water vapour which may condense on the trap or the detector, therefore, temperatures greater than 60°C should not be used (Westendorf, 1981). Practicality should also be taken into consideration since a long purge time would lead to
excessively lengthy runs.

1. Random Centroid Optimization

In order to optimize the purge conditions (sample size, sample temperature and purge time), at least three levels should be used to obtain a proper trend. Changing one factor at a time while keeping the other two constant would result in 27 experiments which may be costly and tedious. Several optimization computer programs are available in which the number of experiments are reduced and trends can be more easily seen (Nakai, 1990).

Random Centroid Optimization, a modification of the Morgan-Deming simplex optimization, offered the following advantages: 1. Simple computer operation, 2. high search efficiency, 3. no need for boundary constraints (limits as to where the search for the optimum can take place) which frequently caused search stalling, 4. less chance on arriving at a local optimum than a global optimum and finally, 5. there is a chance of finding new combinations of the factors (Nakai, 1990).

The Random Centroid Program explores for the optimum conditions in a given search space. This space is defined by the constraints placed on the factors being optimized. The program calculates a random set of experiments which include the centroid points. With the results of these experiments and the mapping process, the search space is narrowed. This
procedure is repeated until the optimum is reached.

G. SENSORY EVALUATION

As a more objective tool than gas chromatography for depicting the aroma and taste of the mango samples, sensory analysis in the form of Quantitative Descriptive Analysis (QDA) was used. QDA utilizes trained individuals to identify and quantify the sensory properties of a product in a statistically reliable manner (Rutledge and Hudson, 1990 and Stone et al., 1974). The first step in this method is the development of the language or descriptive terms to characterize both the aroma and taste of the given product. These terms are derived by the entire panel in a group discussion setting with the panel administrator providing leadership and guidance but not actively participating in the product evaluation (Stone et al., 1974 and Larmond, 1985). The product itself (in this case mangos) were used for both the language development, selection and training of the judges (Stone et al., 1974; Powers, 1988). The samples used for language development should encompass a wide range of treatments including the extremes to help minimize end-order effects (avoidance of extremes) and encourage full use of the scale (Stone and Sidel, 1985). The terms must be well defined and understood by all the judges in the panel and when possible designated an appropriate reference material (Stone and Sidel, 1985).
A questionnaire with specific instructions and a list of these terms is subsequently presented to the judges individually in isolated booths. A 15 cm line with anchor points 1.5 cm from each end is designated for each term - the anchor point describing the full range for that term (Stone et al., 1974). Unlike the Flavour Profile Method, a centre anchor is not used in QDA for two reasons: the first is to allow the judges more freedom in manipulating the scale, avoiding the need to fit their perceptions into designated categories. The second is the need for statistical analysis of the data in which avoidance of a middle anchor would meet the ANOVA assumption of errors being normally distributed (Powers, 1988). In fact, removal of the centre anchor was found to reduce variability by 10-15% (Stone and Sidel, 1985).

Replications of the samples were recommended to improve the reliability of the data and provide information about the consistency of individual judges as well as the entire panel (Stone and Sidel, 1985).

Several researchers have used QDA in conjunction with gas chromatography to characterize the flavour of a product. For example, Aishima et al. in 1987 analyzed and compared strawberry fruit flavour by gas chromatography and verified the results with QDA.

III. MATERIALS AND METHODS

A. SAMPLES
1. **For Optimization of GC Parameters**

Tommy Atkins mangos were purchased from a local wholesaler (Van Whole Produce, Vancouver, B.C.). These mangos were ready for distribution to local outlets and, therefore, judged ready to eat. The wholesaler had received the mangos from Mexico and kept them in cold storage (15 C) until their sale.

2. **For Actual GC Experimentation**

Sample collection of both Tommy Atkins and Haden took place in Tecoman, Colima which is a town located on the Pacific coast of Mexico. Four lots were collected for each cultivar: size 12, 14, 16 and mixed (the size reflected the number of mangos which were able to fit in a prefabricated 5 Kg. box). The first 3 lots (sizes 12, 14 and 16) for both cultivars and both ripening conditions were obtained from Desarrollo Fruticola Del Valle S.P.R. de R.L. (an orchard owned by the Fernandez brothers). The miscellaneous lot (lot 4) for each variety consisted of a mixture of sizes and was comprised of the following: the mangos ripened on trees came from different orchards (Tommy Atkins from Empacada Coliman and Haden from Empacadora de Mango Chulavista both in Tecoman, Colima) and those ripened in storage were bought in Vancouver from Van Whole Produce Ltd. instead of processed in Mexico. Each lot was comprised of 14 samples: 7 ripened on a tree and 7 ripened in storage. Of each of these 7 samples, 5 were
individual mangos and the other 2 consisted of a blend of 10 mangos each (total number of mangos used per lot= 50) All mangos were randomly chosen within their size and method of ripening.

Total number of samples:

2 varieties x 4 lots x 14 samples/lot = 112 samples for analysis

Mangos that were ripened on a tree were harvested at a ready-to-eat stage and no further processing took place except for the sample preparation described below.

Mangos that were to ripened in storage were harvested at the fully mature but green stage (determined by local pickers.)

3. For Sensory Analysis

The following samples from the ones used for G.C. analysis were also utilized for the sensory analysis: Tommy Atkins ripened on tree, Tommy Atkins ripened in storage, Haden ripened on the tree and Haden ripened in storage. The second mix of 10 mangos for all four lots (sizes 12,14,16 and mixed) was tested for a total of 16 different samples.

In addition, to train the sensory panel two other cultivars of mango were obtained from Mexico: Irwin and Manila. Both were ripened on the tree.

4. For Gas Chromatography/Mass-Spectroscopy

GC/MS was used to identify the compounds present in four
representative samples (Tommy Atkins ripened on the tree and ripened in storage, and Haden ripened on the tree and ripened in storage). The first mixed sample of lot 2 (size 14) for each of the four conditions was used for the analysis.

B. SAMPLE PREPARATION FOR ANALYSIS

1. Samples ripened in storage

These mangos underwent the following processing steps:

**Pre-storage**

**Step 1: washing**

Fruit was soaked in water at ambient temperature with detergent added. They were then passed under brush rollers and sponges which scrubbed the fruit clean.

**Step 2: drying**

Mangos were dried in a current of hot air.

**Step 3: classification by size**

A specially designed machine separated the fruit according to size.

**Step 4: hot water treatment**

Mangos were soaked in a bath at 47.5 C as required by the USDA to prevent anthracnose disease. The time of submersion of the fruit depended on its size (and therefore, its weight). All fruits in this experiment were held in the hot water for 75 min.

**Step 5: cold water treatment**
Immediately after the mangos were removed from the hot water bath, they were immersed in cold water (21-25 C) for 20 min to cool the fruit and, thereby, minimize heat damage.

**Step 6: storage**

Mangos were then classified by hand according to size, weighed individually on a table-top electronic balance, and placed in specially designed boxes which allow air to circulate. The boxes were then stacked on racks and stored at a temperature of 15 C until fully ripe.

Mangos were judged to be fully ripe by the local owner after 10 days of storage (Aina, 1990). Judgement was based mainly on the softness of the fruit as it yielded to hand pressure.

**Post-storage**

1) Mangos were peeled, then cut and the pit removed.

2) A puree was made with a Waring blender and then passed through a mesh strainer (2.0 mm openings) to remove the fibrous material. For the mixed samples, the puree of 10 mangos was blended thoroughly. The pulp was used since El-Samahy et al. in 1982 pointed out that the majority of the flavouring compounds were present in the pulp portion.

3) The puree was then placed in 40 mL vials with an appropriate code and frozen at -20 C. The mango pulp may be stored from 3 to 6 months at -20 C without the need for heat
treatment (Pino et al., 1989 and Sakho et al., 1985).

Samples were stored in dry ice and then transported from Mexico to Canada by plane and immediately placed in -20 C at arrival.

2. Gas chromatography

Optimization

The Tommy Atkins samples purchased in Vancouver were defrosted and the appropriate weight of the sample measured into a test tube. The sample was then diluted with distilled, deionized water to 10.0 g. The analysis of volatiles was then carried out.

The SRI Gas Chromatograph:

Recently SRI in Torrance, California, developed a compact, low cost and portable gas chromatograph equipped with a purge and trap and an FID detector (8610 model) which was suitable for routine analysis (Figure 1). The purge and trap unit was heated with a temperature controlled hot wand. The traps used were supplied by SRI (Torrance, CA) and when cryo-trapping was utilized, the liquid CO$_2$ was acquired from Medigas Pacific Ltd., and connected directly to the GC.

The hydrogen and helium gases used for the detector and the nitrogen purge gas were of ultra high purity and acquired from Medigas Pacific Ltd., Vancouver. Gas filters and purifiers (Supelco chromatography products) were installed in all gas lines between the gas cylinders and the chromatograph to
Figure 1. The SRI (Model 8610) gas chromatograph
remove water, oxygen and oil from the gas. An air compressor and filter were built into the GC so no air tank was required. The nitrogen gas flow rate was 10 mL/sec and the helium (carrier gas) flow rate was 3.34 mL/min in the capillary column and 4.62 mL/min for the megabore column.

The flow rates were obtained from the retention time of methane after injection at 45 C. The formula used to calculate the flow rate was:

\[
\text{Flow rate (mL/min)} = \frac{\pi r^2 L}{t_m}
\]

where

- \( r \) = column radius in cm
- \( L \) = column length in cm
- \( t_m \) = retention time of methane in min

To operate the GC and collect and analyze the data, the GC was linked to an IBM compatible 386K personal computer loaded with Peaksimple II, a software package provided by SRI. This program contains many features including various options for the integration of the resultant peaks.

2) Routine analysis

The appropriate sample vial was defrosted under running cold water. As the optimization results dictated, 0.11 g sample was weighed out into a test tube and diluted with distilled, deionized water to 10.0 g. The internal standard was added (20 microliters of diluted standard) and the contents mixed thoroughly. The analysis was then carried out.
using the optimized conditions and the GC parameters outlined in the section on optimization.

The needles used for the internal standard were cleaned with purge and trap grade methanol and distilled water between applications and then left to dry in a drying oven.

3. **Samples ripened on the tree and those used for the optimization of the GC parameters**

   There was no pre-harvest control of these samples and the post-harvest treatment was the same as for those mangos ripened in storage.

4. **Gas chromatography/ mass spectroscopy**

   The conditions for the GC/MS analysis were the following: A Hewlett Packard Model 5840 GC was equipped with a Tenax trap without cryotrapping. A 30 m X 0.53 mmI.D. DB-624 megabore column with 3.0 µ film thickness was temperature programmed to increase by 8 deg/min from 30 C to 225 C. The optimized purge conditions were used with the mass spectrometer except the sample was not diluted since water accumulation became a problem. Instead, 6 g of the sample were heated to 51 C in a water bath and purged for 8 min.

   The mass spectrometer was a Hewlett Packard Model 5985B Quadropole with a Teknibent data system operated with the following parameters:
   
   - Ion source temperature: 200 C
ionization voltage: 70 eV
interface temperature: 250 C
scan rate: 40-350 AMU at 0.5 sec.

5. **Titratable acidity and pH**

The appropriate sample was defrosted in cold water and 10.0 g weighed out. The sample was then diluted with the addition of 25.0 mL of distilled, deionized water. The pH was measured with a Fisher pH/ion meter (Model 420) as the sample was being stirred. It was then titrated to a pH of 8.1 with 0.10 N NaOH. Titratable acidity was calculated as % citric acid per gram sample.

6. **Sugar analysis**

The appropriate sample vial was defrosted with cold water. A 1.0 g aliquot was weighed out and diluted with distilled, deionized water in a 10.0 mL volumetric flask. The sample was centrifuged for 15 min at 10,000 RPM in a Sorvall centrifuge (Model RC2-B). For the analysis, 1.0 mL of the supernatant was diluted again with distilled, deionized water in a 10 mL volumetric flask.

A Boehringer and Mannheim (Biochemica) sucrose/D-glucose/D-fructose enzyme kit was subsequently utilized to analyze the sample. In this ultra-violet method the Varian Carry 210 spectrophotometer was set at a wavelength of 340nm. The analysis is based on determining the glucose content
before and after the hydrolysis of sucrose into glucose and fructose as well as the determination of fructose separately. The results of the separate sugars were then added to obtain the total sugar content of the sample.

7. Sensory analysis
   a) Odour analysis

   Mango samples were defrosted and subsequently heated to 51°C in a hot water bath to simulate the purging temperature optimized for the GC analysis.

   Training of panel members: Development of the Descriptive Terms

   Ten judges (4 males and 6 females) from the Food Science Department at the University of British Columbia gathered in a round-table discussion session to determine the appropriate terms that should be used to describe each sample. Each judge was given the test samples and asked to choose as many descriptive terms as he/she could to characterize the sample’s aroma. The test samples included heated mango puree of Haden and Tommy Atkins both ripened on the tree and in storage as well as heated samples of Irwin and Manila mangos. The resulting terms were then compiled by the panel leader and discussed with all the judges until a clear set of terms was developed and all the judges agreed on their meaning. To aid the judges in their task, standards representing the
descriptive terms were provided.

**Training of panel members: Practice Sessions**

Panelists were introduced to the questionnaire format and asked to judge several samples individually. The results of the tests were discussed amongst all the panelists and fine-tuning of the terms took place.

**Questionnaire Format:**

Each judge was given a set of specific instructions detailing how he/she should place a vertical line across the horizontal lines at the point which best reflected the magnitude of the perceived intensity of that particular attribute. The horizontal line was 15cm long with anchor points 1.5cm from each end and a descriptive term at each of these anchors. The descriptive terms were those determined during the training sessions. An example of a questionnaire is shown in Figure 2. Results were transformed into usable data using SigaScan (Jandel Scientific). This program transforms a point into a numerical value after the endpoints of a line and its intervals are defined.

b) Taste Analysis

In order to verify the validity of the sugar/acid ratio determined by instrumental means, panelists were asked to taste the defrosted mango puree and determine its degree of sweetness and sourness using the same questionnaire format as
Questionnaire for Descriptive Analysis of Mangos

Please evaluate the odour of the coded samples by making a vertical line on the horizontal line to indicate your rating of the particular attribute.

sample code # ______

overall intensity

very weak

very strong

tea-like

very weak

very strong

weed-like

very weak

very strong

fruity

very weak

very strong

pineapple/banana

very weak

very strong

honeydew melon

very weak

very strong

sweet

very weak

very strong

Figure 2. Questionnaire for aroma determination of mango samples using QDA.
for the odour analysis (Figure 3).

Prior to the evaluation of the mango samples, Magnitude Estimation was used to assess judges' perception of sweetness and sourness. Four concentrations of citric acid (0, 0.01, 0.03 and 0.1 percent) and sucrose (0, 0.15, 0.55 and 1.0 percent) were used independently. Testing was performed in individual booths and samples were served at room temperature. Samples were encoded with a three digit code and tested in random order. The panellists were asked to rank the 4 samples in descending order of intensity (ie most sour to least sour and most sweet to least sweet).

Data Collection:

Both taste and odour were analyzed in the same session. The 32 samples (16 in duplicates) were presented to each judge in a random order. Four samples were evaluated per session. Judges evaluated heated samples in individual booths with red light to avoid errors due to colour differences among the samples.

8. Magness Taylor pressure determination

A Magness Taylor penetrometer (Hoskin Scientific Ltd., Vancouver, B.C.) fitted with a 5/16" diameter plunger was used to test the firmness of the fruit. The plunger was inserted with even pressure into the tissue in the plump side of each mango. The reading was made in kilograms.
Sweetness and Sourness

Please taste these samples and evaluate them for their degree of sweetness and sourness. Place a vertical line on the horizontal line to indicate the intensity you perceive.

Sample # __________

[Horizontal line with not sweet on the left and very sweet on the right]

[Horizontal line with not sour on the left and very sour on the right]

Figure 3. Questionnaire for sweet/sour taste determination of mango using QDA
IV. RESULTS AND DISCUSSION

A. GAS CHROMATOGRAPHY

1. Optimization: preliminary

Before any computer optimization could be utilized, preliminary work had to be done to establish some general parameters for the gas chromatograph. The sampling procedure was chosen to be dynamic headspace analysis for the reasons discussed earlier. The column, trap, and temperature program needed to be decided upon for the mango sample using the SRI chromatograph.

Temperature program

A conservative temperature program was used at first to obtain a general perspective of the chromatograms. The temperature program began with holding the oven at 35°C for 10 min, then the temperature was increased at 4 deg/min until 210°C was reached and then holding the temperature at 210°C for 5 min. Modifications of this program were introduced after other parameters were established.

Columns

Three columns were used (all acquired from J&W Scientific, Rancho Cordova, CA). The first was a DB624 capillary column which was 30 m long, had a 0.32 mm diameter and a film thickness of 1.8 µ. The stationary phase of this medium polarity column was comprised of cyanopropyl, phenyl, dimethyl polysiloxane crosslinked and bonded to fused silica. The Megabore counterpart of the DB-624 (30 m long, diameter of
0.53mm and film thickness of 5μ) was the second column used. Finally, a highly polar capillary column DB-WAX (30 m long, 0.32 mm diameter and 0.5μ film thickness) with a stationary phase of polyethylene glycol crosslinked and bonded to fused silica was tested. This column is highly desirable for purge and trap when separating low boiling, polar molecules (Takeoka and Jennings, 1984). A non-polar column was not used since mango aroma consists of complex mixture of polar and nonpolar compounds.

**Traps**

Charcoal and Tenax traps were compared without cryo-trapping at first. When liquid CO₂ is used to cool the trap, there is a greater chance for water accumulation when the trap is thermally desorbed. This extra moisture can easily extinguish the FID. Later cryo-trapping proved to be a greater benefit than a problem.

A dry purge (an empty test tube attached to the purging apparatus) of 19 min followed immediately after the purging of the sample so that excess water accumulated during purging was eliminated (Westendorf, 1981). The traps were initially baked for 15 min at 230 C (Tenax) and 450 C (charcoal) to eliminate any retained volatiles, a standard procedure in most analyses (Westendorf, 1981). However, this method did not prove sufficient for the mango samples. Figure 4 shows a blank run (10.0 mL distilled water) after a 15 min baking period at 230 C. The eluted peaks were probably due to the
Figure 4. GC chromatogram of a blank run (10 mL distilled water) of a Tenax trap after 15 min of baking at 230°C.
trap, necessitating longer baking. Figure 5 shows the same blank run after 30 min of baking at 230 C. Therefore, the baking time was increased to 30 min.

Purge conditions were to be optimized by Random Centroid Optimization but the upper and lower limits of the factors had to be established. Therefore, a sample size of 5.0 grams was arbitrarily chosen as the starting point.

Figure 6 shows a chromatogram of the 5.0 g of Tommy Atkins sample (diluted to 10.0 mL with distilled water) and heated to 45 C for 15 min. The column used was the capillary DB624 with a Tenax trap. Immediately it was noted that there was a substantial overloading of the column (huge peak at the right hand side of the chromatogram). This overloading was highly undesirable since the one huge peak may be masking other peaks. There were various ways to remedy this problem and the first was to reduce the sample size thereby reducing the quantity of the volatiles. Figures 7-9 show chromatograms with the same conditions as in Figure 6 but reduced sample size (2.5g, 0.5g and 0.1 g respectively). Although the overloading of the column decreased substantially, reduction and elimination of other peaks also occurred. To test the effect of the stationary phase on the peak separation, a DB-WAX column was installed. A 0.25g mango sample was heated to 40 C for 15 min. The resultant chromatogram shown in Figure 10 indicated a decrease in the resolution of the peaks as well as increasing the baseline. Both these factors can be
Figure 5. GC chromatogram of a blank run (10 mL distilled water) of a Tenax trap after 30 min of baking at 230 C.
Figure 6. GC chromatogram of volatiles from five grams sample of Tommy Atkins heated to 45 °C for 15 min (dry purge 19 min), concentrated on a Tenax trap and eluted onto a capillary DB-624 column.
Figure 7. GC chromatogram of volatiles from a Tommy Atkins sample (2.5 g) with the same conditions as in Figure 6.
Figure 8. GC chromatogram of volatiles from a Tommy Atkins sample (0.50 g) with the same conditions as in Figure 6.
Figure 9. GC chromatogram of volatiles from a Tommy Atkins sample (0.10 g) with the same conditions as in Figure 6.
Figure 10. GC chromatogram of volatiles from a Tommy Atkins sample (0.25 g) heated to 40 C for 15 min, concentrated on a Tenax trap and eluted onto a capillary DB-WAX column.
attributed to the instability of the stationary phase (Takeoka and Jennings, 1984). In an attempt to improve the resolution and separation of the peaks the following changes were made and the results shown in Figure 11: lowering the carrier gas flow rate (1.96 mL/min), decreasing the purging temperature (28 C) and using a smaller sample size (0.1 g). A slight improvement was noted but overloading (even at such small sample size and no heating during purging) was apparent. Figure 12 shows a chromatogram obtained using the same conditions as in Figure 11 except a charcoal instead of a Tenax trap was used to concentrate the volatiles. No substantial improvement resulted from this change and, therefore, the capillary DB-624 was judged better than the capillary DB-WAX column.

To try and further improve the resolution, a Megabore column (DB-624) was installed. The helium flow rate was increased to 4.62 mL/min since higher flow rates are required for larger diameter columns (to keep the head-pressure constant). Again a 0.5 g sample was used and it was purged for 12 min at 39 C. The temperature program was also modified slightly to ramp at 2 deg/min instead of 4 deg/min. The results are shown in Figure 13 (charcoal trap) and Figure 14 (Tenax trap). The Tenax trap resulted in a better resolution of the first clump of peaks and was regarded as superior to the charcoal trap. Westendorf in 1981 pointed out that if a compound that is normally trapped on Tenax is carried into the
Figure 11. GC chromatogram of volatiles from a Tommy Atkins sample (0.10 g) purged at 28 C for 15 min, concentrated on a Tenax trap and desorbed onto a capillary DB-WAX column with a helium flow rate of 1.96 mL/min.
Figure 12. GC chromatogram of volatiles from a Tommy Atkins sample with the same conditions as in Figure 11 except volatiles were concentrated on a charcoal trap.
Figure 13. GC chromatogram of volatiles from a Tommy Atkins sample (0.50 g) was purged at 39 °C for 12 min, concentrated on a charcoal trap and desorbed onto a Megabore DB-624 column with a helium flow rate of 4.62 mL/min. The temperature program was modified to ramp at 2 °C/min instead 4°C/min.
Figure 14. GC chromatogram of volatiles from a Tommy Atkins sample with the same conditions as in Figure 13 except sample was desorbed onto a Tenax trap.
charcoal matrix, broader peaks and even some loss in recovery may occur before the breakthrough volume is reached resulting in inferior chromatograms. Therefore, Tenax was preferred to charcoal as the adsorbent for trapping the mango volatiles.

Subsequently, the use of liquid CO$_2$ to lower the temperature of the Tenax trap was tested. As expected, the lower boiling point volatiles were better retained than without the cryo-trapping. Since the liquid CO$_2$ was used in bursts, not enough water accumulated to extinguish the FID. In addition, dry purging was no longer required saving 19 min on the run time. Therefore, for the Random Centroid Optimization, cryo-trapping was used.

The capillary column, having a more linear baseline, was preferred to the Megabore column. The larger capacity Megabore column tended to exhibit a rising baseline due to its much thicker stationary phase which leads to increased bleeding. In addition, the capillary column is more frequently used by researchers enabling comparison with existing literature.

2. Optimization: Random Centroid Optimization (RCO)

With the basic parameters established, the purge conditions were then optimized using RCO. The preliminary optimization narrowed down the limits for the sample size, purge temperature and purge time. The search space for the optimum, therefore, was confined to sample size ranging from 0.05 g to 0.25 g; purge temperature varying from 35 C to 55 C
and purge time ranging from 5 min to 15 min.

In RCO, a response factor was used to numerically compare the results of the prescribed experimental conditions. In gas chromatography the Chromatographic Response Function (CRF) has been used to quantitatively represent a chromatogram (Aishima et al., 1990). The CRF takes into account both the number of peaks and their resolution and was calculated with the equation below:

\[
\text{CRF} = n + \sum_{i=1}^{n} P_i
\]

where \( n \) is equal to the number of peaks in the chromatogram and \( P_i \) was equal to \( f \) divided by \( g \) as demonstrated in Figure 15. The larger the CRF value the greater the number of peaks and the better the resolution between them.

The first nine experiments outlined in Table 1 were the result of the first cycle of the RCO program. The vertex numbers were provided by the program, while the actual numbers represent the order of experimental execution that resulted from manual random selectivity. The outcome of these experiments as measured by the CRF were used to generate the centroid experiments (Table 1). All this data was mapped to aid in narrowing the search space for the optimum (Figure 16). The mapping process aids in visualization of the experimental response surface indicating the trend of the data (Nakai, et al., 1984). The mapping process, therefore, helped confine the search space for the second cycle to a sample size of
Figure 15. The resolution between two peaks was calculated by dividing the value of $f$ by $g$. The resulting $P_1$ value contributed to the Chromatographic Response Function (CRF).
<table>
<thead>
<tr>
<th>Vertex number</th>
<th>Actual number</th>
<th>Sample size</th>
<th>Sample temp.</th>
<th>Purge time</th>
<th>CRF**</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>0.20</td>
<td>40</td>
<td>9.44</td>
<td>12.46</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>0.11</td>
<td>50</td>
<td>6.24</td>
<td>36.84</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>0.10</td>
<td>37</td>
<td>11.36</td>
<td>19.25</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0.22</td>
<td>40</td>
<td>8.59</td>
<td>19.03</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>0.08</td>
<td>48</td>
<td>6.99</td>
<td>26.42</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>0.12</td>
<td>52</td>
<td>10.56</td>
<td>23.37</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.25</td>
<td>50</td>
<td>5.45</td>
<td>21.04</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>0.08</td>
<td>49</td>
<td>12.33</td>
<td>23.97</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>0.21</td>
<td>40</td>
<td>14.24</td>
<td>19.00</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>0.09</td>
<td>49</td>
<td>8.52</td>
<td>25.48</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>0.09</td>
<td>49</td>
<td>7.25</td>
<td>24.18</td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>0.12</td>
<td>51</td>
<td>5.30</td>
<td>29.18</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>0.10</td>
<td>50</td>
<td>6.31</td>
<td>26.12</td>
</tr>
<tr>
<td>14</td>
<td>16</td>
<td>0.18</td>
<td>48</td>
<td>8.94</td>
<td>22.43</td>
</tr>
<tr>
<td>15</td>
<td>12</td>
<td>0.17</td>
<td>48</td>
<td>6.90</td>
<td>22.33</td>
</tr>
<tr>
<td>16</td>
<td>14</td>
<td>0.09</td>
<td>51</td>
<td>9.55</td>
<td>26.25</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>0.10</td>
<td>50</td>
<td>7.05</td>
<td>17.72</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>0.10</td>
<td>51</td>
<td>7.30</td>
<td>32.25</td>
</tr>
</tbody>
</table>

* random selection of vertex numbers  
** Chromatographic Response Function

Table 1. Experimental plan generated by Random Centroid Optimization (RCO). The vertex number represents the RCO order while the actual number was obtained by randomizing the vertex numbers. The latter was used to carry out the experiments.
Figure 16. Mapping results of the first cycle of experiments generated by RCO. Figure A, B and C outline the trends for sample temperature, sample size and purge time, respectively.
0.08 g to 0.18 g, sample temperature ranging from 47 C to 52 C and a purge time varying from 5 min to 10 min. Using these limits, the same procedure was followed as in cycle 1 and the optimum was reached after 18 experiments. Results of all the experiments are found in Table 1. Although Vertex 2 has the largest CRF value, in actuality the chromatogram was less acceptable than the one obtained in Vertex 18 due to inferior peak resolution. The larger number of peaks obtained using condition in Vertex 2 contributed heavily to the resultant CRF value. This example demonstrates the importance of subjective evaluation to arrive at the best conditions. Subsequent mapping of all the results indicated the optimum conditions for the purge parameters (Figure 17). Therefore, using the RCO approach, best results were obtained when a 0.11 g. mango sample was purged at 51 C for 8 min. Figure 18 compares a chromatogram obtained using Vertex 3 conditions of Table 1 and the optimum conditions. A substantial improvement was apparent in the resolution of the peaks as well as the lowering of the baseline. The RCO program eliminated the problem with overloading of the later eluting peaks while retaining the minor ones. Although the resolution of the peaks does not seem appropriate for a capillary column, limitations inherent to this GC model play an important role. The oven may not be properly insulated to obtain satisfactory temperature control leading to trailing peaks and some band broadening. However, for the current application it was
Figure 17. Mapping results of all 18 experiments generated by RCO. Figure A, B and C outline the trends for sample temperature, sample size, and purge time, respectively. The arrow indicates the predicted optimum conditions for each factor.
Figure 18. Reproductions of Tommy Atkins mango chromatograms obtained using the RCO program. The dotted line chromatogram resulted from the conditions outlined in vertex 3, Table 1, while the solid line chromatogram was obtained after optimization of the GC parameters with RCO.
judged satisfactory.

3. **Internal Standard**

An internal standard was required to quantitate the gas chromatography peak areas. The compound used as the internal standard should not be part of the mango effluent as well as elute in an area where no other volatiles are present. Previous mango researchers have used tridecane and 1-heptanol with Alphonso and Baladi varieties (Engel and Tressl, 1983), ethyl acetate with Tommy Atkins and Kitts varieties (MacLeod and Snyder, 1985), alloocimene, methylphenylacetate and undecane-1-ol with the Alphonso variety (Idstein and Schreier, 1985) and heptan-1-ol for Corazon, Bizcochuelo and Super Haden varieties (Pino et al., 1989). Following these authors' results, several alcohols and hydrocarbons were tried: 2-methyl-1-propanol, 1-pentanol, 3-methyl-1-butanol, n-octadecane and n-dodecane (all acquired from PolyScience Corp., Niles, Illinois), however, none gave adequate results. The problem may have been due to the method of analysis (headspace vs. distillation and extraction) as well as varietal differences. A ketone, 4-methyl-2-pentanone, was finally chosen since it showed good recovery with the purge and trap system and it eluted in an area where no other peaks were found. To normalize the data, all the peak areas from the eluted volatiles were divided by the peak area of the standard.
Normalized peak area = \[
\frac{\text{area of peak}}{\text{area of standard}}
\]

This method allowed the comparison of the different chromatograms for all the samples (Mohler Smith, 1990).

4. **Repeatability of Chromatograms**

Three chromatograms were run with the optimum conditions obtained from the RCO to determine the repeatability of the gas chromatograph. The internal standard was added at a concentration of 0.1 ppm after the sample was diluted and before beginning the purge sequence. The mango sample was the same as used for the optimization procedure. Five major peaks were compared for their repeatability after they had been normalized with the internal standard as discussed above. Figure 19 compares the three chromatograms and Table 2 shows the mean, standard deviation and the coefficient of variation obtained by dividing the standard deviation by the mean. These results indicate that the G.C. showed good repeatability for these samples.

5. **Comparison of peaks and their selection for multivariate analysis**

Peaks were compared by their retention times— that is a peak in one chromatogram that had the same retention time (given a narrow range) as a peak in another chromatogram were said to be the same compound. The range for the retention
Figure 19. Repeatability of three chromatograms of Tommy Atkins using the optimized conditions.
Table 2. The average, standard deviation and coefficient of variation of 5 peaks obtained from each of three Tommy Atkins chromatograms run under optimized conditions to test the repeatability of the SRI GC.
time was chosen by site and careful comparison of all the chromatograms.

A total of twenty eight different major and minor volatiles were detected in both varieties in differing quantities and frequency.

Of these twenty eight, seven were chosen for the multivariate analysis processing. The selection was based on the degree of area fluctuation exhibited by each volatile—those volatiles whose area changed the most from sample to sample were chosen for the multivariate analysis. The average areas for the seven peaks ranged from 0.15 to 0.87 with standard deviations ranging from 0.06 to 0.76. The difference between the samples within the lot, especially for the ones ripened on the tree, contributed to the large standard deviation.

Figures 20 and 21 show sample chromatograms of Tommy Atkins and Haden ripened on the tree and ripened in storage, respectively. Few peaks were seen due to the sampling procedure being purge and trap as noted by Bartley and Schwede in 1987.

6. Effect of fruit size

The lot numbers assigned to each sample designated their size (ie lot 1= size 12, lot 2= 14, lot 3= 16 and lot 4 was a mix). The four lots showed similar chromatograms for each variety and ripening condition including lot 4 which came from
Figure 20. Sample chromatogram of Tommy Atkins mango ripened on the tree.
Figure 21. Sample chromatogram of Haden mango ripened in storage.
different orchards. As an example, Figure 22 shows lots 1 and 3 of the Haden variety ripened in storage. Peak 1 of lot 1 volatiles has a slightly larger area than the same peak in lot 3 but otherwise the two lots were very alike. Similarly, lots 2 and 4 of Tommy Atkins ripened in storage show comparable peak distribution (Figure 23). Individual number three in lot 4 shows no peak 2, but it is an exception (probably an outlier) and can’t really be compared with lot 2 since it differed from other samples within its own lot. Suryaprakasa Rao et al. in 1972 also noted that the size of Baneshan mangos did not affect the quality of the ripe fruit nor the rate of ripening. In addition, Soule Jr. and Harding in 1956 found little effect of size of the mango on the chemical parameters studied. Therefore, the mango size was not considered as a suitable means of classification and the different lots were used interchangeably in the computer analysis.

B. GAS CHROMATOGRAPHY/ MASS SPECTROMETERY

Figures 24 and 25 show the identification of the major peaks in Tommy Atkins and Haden mangos. The three major compounds identified were \( \alpha \)-Pinene, Car-3-ene and \( \beta \)-Phellandrene (their structure and mass spectrum are shown in Figure 26). All three compounds have a molecular weight of 136 and a general molecular formula of \( \text{C}_{10}\text{H}_{16} \).

Car-3-ene was found to be the most abundant volatile in Tommy Atkins from Florida (MacLeod and Snyder, 1985) and from
Figure 22. Simulated chromatograms of lot 1 (size 12) = A and lot 3 (size 16) = B, of Haden mangos utilizing adjusted areas of the seven peaks used for multivariate analysis.
Figure 23. Simulated chromatograms of lot 2 (size 14) = A, and lot 4 (mixed sizes) = B, of Tommy Atkins mangos utilizing adjusted areas of the seven peaks used for multivariate analysis.
Figure 24. Mass Spectroscopy identification of major peaks of Tommy Atkins sample.

Peak 1 = unknown
Peak 2 = unknown
Peak 4 = unknown
Peak 5 = standard (4-methyl-2-pentanone)
Peak 5 = α-Pinene
Peak 6 = β-Phellandrene
Peak 7 = Car-3-ene
Peak 1 = unknown  
Peak 5 = α-Pinene  
Peak 6 = β-Phellandrene  
Peak 7 = Car-3-ene

Figure 25. Mass Spectroscopy identification of major peaks of Haden sample.
Figure 26. Mass Spectrometer spectrums and chemical structures of \( \alpha \)-Pinene, Car-3-ene and \( \beta \)-Phellandrene.
Brazil (Franco et al., 1991), in Cuban mangos (Pino et al., 1989), in Venezuelan mangos (MacLeod and Pieris, 1984) as well as in Haden from Brazil (Franco et al., 1991). This compound was also found in lesser quantities in African mangos (Sakho et al., 1985) and in the Indian variety of Alphonso (Idstein and Schreier, 1985). Car-3-ene was found to have a 'mango leaves' aroma and, therefore, considered highly desirable in mangos (MacLeod and Snyder, 1985). α-Pinene and β-Phellandrene were also found in most of these varieties, but in varying quantities. The first couple of peaks could not be identified probably due to the lack of cryo-trapping in the GC/MS system (i.e. the Tenax trap by itself failed to retain these highly volatile compounds).

C. MULTIVARIATE ANALYSIS TECHNIQUES

In an attempt to classify the mango samples into two cultivars and two ripening conditions, three multivariate analysis techniques were used. These will be briefly discussed below. All statistics and graphing were performed by SYSTAT (Wilkinson, 1989) and Lotus1-2-3 version 3.1.

1. Principal Component Analysis (PCA)

PCA is one of the most commonly used multivariate analysis techniques in food science. PCA aims at reducing the number of variables required for data manipulation, and in some cases it was also used for classification purposes. PCA
proceeds by searching for linear combinations of variables which account for the maximum possible proportion of variance in the original data (Piggott, 1986).

The first step of the analysis is the orthogonal rotation of the N axes of the original pattern space. The N axes can then be expressed in terms of the original axes. Each new axis is expressed in the following linear function:

\[ PC_1 = a_{11}X_1 + \ldots + a_{n1}X_n \]

where:

- \( PC \) = principal component (N)
- \( a \) = constants of transformation (loadings)
- \( X \) = vectors defining the location of the variable points in object space
- \( i = (1,N) \)
- \( n = \) number of variables

The first principal component needs to explain the maximum variance, the second principal component which is orthogonal to the first must explain the largest amount of the residual variance etc. (Derde and Massart, 1985 and Powers and Ware, 1986). The PCA loadings indicate the importance of the original parameters in the direction of the principal component (Derde and Massart, 1985).

2. Linear Discriminant Analysis (LDA)

LDA is a supervised technique in which mutually exclusive groups are obtained from a predetermined classification. As stated by Powers in 1986, the basic concept of LDA is to find,
through a transformation of the original data, a linear combination of variables so that the mean distance between classes can be maximized. In some cases, it is advantageous to use Principal Component Analysis first to reduce the number of variables and subsequently apply LDA (Mohler Smith and Nakai, 1990). The basic equation for the discriminant function was the following:

\[ L = b_1 x_1 + b_2 x_2 + \ldots + b_p x_p \]

where:

- \( L \) = linear composite score
- \( b \) = the weight of the variable

The goal, therefore, is to obtain the set of \( b \)'s that will maximize the between group variability while minimizing the within group differences. The \( b \)'s are selected in such a way that \( b_1 \) reflects the largest group difference, \( b_2 \) accounts for the largest group difference not accounted for in \( b_1 \) etc. Ideally, the first few functions are adequate to express almost all the vital group differences (Manly, 1986 and Powers and Ware, 1986).

Once the discriminant function has been established, a need arises to determine its success in classification of the groups. The computer program utilized in this work, SYSTAT (Wilkinson, 1989), used the Mahalanobis distance from the observation to the group centre in order to classify the samples into their respective groups.

3. Principal Component Analysis—Similarity
A new classification method was developed in our laboratory in which a sample is compared to a predefined standard. Principal Component Analysis-Similarity is a nonsupervised pattern recognition technique that uses mult-criteria to discriminate between desired and undesired characteristics.

IBM BASIC was used to write a PCA-SIM (Appendix 1) which consisted of five parts (Figure 27). Principal component (factor) scores were obtained by any appropriate statistical program (i.e., Systat, SAS) and used in the PCA-SIM program. This step was critical since PCA reduces the number of variables required to be analyzed. The number of factor scores used in the analysis may be determined in several ways.

The first, and most common, is by choosing those factors whose eigenvalues are equal to or greater than one. This elimination criteria is based on the idea that a component having an eigenvalue of one is explaining the average proportion in the data so a value greater than one explains a disproportionately large proportion of variance.

A second method is that of cross-validation and it includes performing the analysis several times with different subsets of objects omitted and estimating these omitted scores from the model. Once the number of scores has been chosen, a new file is created which contains just the relevant factors and this becomes the working file which can be
Figure 27. Flow diagram of Principal Component Analysis—Similarity (PCA-SIM) program.
used with the new program (Powers and Ware, 1986). For the present analysis, the eigenvalue method was chosen as in related literature (Cortoneo et al., 1990, Mohler Smith and Nakai, 1990).

A reference or standard needs to be selected. This reference is one of the samples obtained or some ideal sample to which all other data would be compared. An arbitrary magnification coefficient was subsequently chosen. This coefficient (which is randomly selected) should usually be between the magnitudes of 10-30. It aids in magnifying the difference between the regression lines, elucidating clearly the differences/similarities between the samples or between a sample and the reference.

The corrected principal component scores \( V'_i \) are obtained for each sample by using the following equation:

\[
V'_i = (S_i - R_i)M + V_i
\]

where:

- \( V'_i \) = corrected PCS
- \( S_i \) = principal component scores of the sample
- \( R_i \) = principal component scores of the reference
- \( M \) = an arbitrary magnification value to obtain visually clear results
- \( V_i \) = an adjusted variance obtained from PCA for a particular principal component
Linear regression is performed on the corrected PCS of the reference against each of the samples from which a slope (S) and a coefficient of determination (R) are obtained. These values are then plotted (i.e. S vs. R) where the relative similarity of each of the samples to the standard could be visualized. If a particular sample was of interest, its corrected PCS could then be plotted against those of the standard. A perfect match would be obtained if the points were to lie on a 45 degree line, therefore, any variation was compared to such a line. The differences between the corrected PCS of the reference and sample can then be accounted for by referring to the factor loadings of PCA.

D. MULTIVARIATE ANALYSIS OF GAS CHROMATOGRAPHY RESULTS

1. Principal Component Analysis

PCA was performed on the seven normalized areas of all 112 mango samples. The first three principal components had an eigenvalue greater than one and accounted for 78% of the total variance (Table 3a). The factor loading for each principal component (factors) are shown in Table 3b- the higher the absolute value of the loadings the more that component (peak area) contributed to the variance between the samples. A 3-dimensional plot of a representative lot from each group (i.e. Tommy Atkins ripened on the tree,"A", and ripened in storage, "B" and Haden ripened on the tree,"C", and ripened in storage,"D") is shown in Figure 28. Seven points
### Table 3. Results of PCA (a= eigenvalues and % area explained and b= factor scores of peak areas) of the seven normalized peak obtained from GC chromatograms for all samples.

**a**

<table>
<thead>
<tr>
<th>Factor #</th>
<th>Eigenvalues</th>
<th>% var. explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.852</td>
<td>40.74</td>
</tr>
<tr>
<td>2</td>
<td>1.515</td>
<td>21.64</td>
</tr>
<tr>
<td>3</td>
<td>1.096</td>
<td>15.65</td>
</tr>
</tbody>
</table>

**b**

**COMPONENT LOADINGS**

<table>
<thead>
<tr>
<th>Peak #</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.850</td>
<td>0.180</td>
<td>0.074</td>
</tr>
<tr>
<td>2</td>
<td>0.845</td>
<td>-0.391</td>
<td>-0.214</td>
</tr>
<tr>
<td>5</td>
<td>0.787</td>
<td>0.013</td>
<td>0.192</td>
</tr>
<tr>
<td>4</td>
<td>0.698</td>
<td>0.496</td>
<td>-0.304</td>
</tr>
<tr>
<td>6</td>
<td>0.520</td>
<td>0.044</td>
<td>0.430</td>
</tr>
<tr>
<td>3</td>
<td>0.105</td>
<td>-0.736</td>
<td>0.582</td>
</tr>
<tr>
<td>1</td>
<td>-0.169</td>
<td>0.736</td>
<td>0.626</td>
</tr>
</tbody>
</table>
Figure 28. Plot of PCA factor scores (Principal Components), derived from GC results for four representative lots: A= Tommy Atkins ripened on the tree, B= Tommy Atkins ripened in storage, C= Haden ripened on the tree and D= Haden ripened in storage.
representing the individual and mixed sample were plotted for each of the four groups. This plot, as in most 3-dimensional graphs, is difficult to visualize. Although the Haden mango samples (C and D) are reasonably well separated into their groups, the Tommy Atkins samples (A and B) could not be well differentiated. Similar results were obtained when using various combinations of different lots. Therefore, no clear classification was noted using the PCA program.

2. Linear Discriminant Analysis

LDA was performed on the seven normalized peak areas for the same four lots (28 samples) used in PCA and compared to LDA performed on the three factor scores obtained from first running PCA. Table 4 shows the results of the predicted values for both the principal component and individual peak area methods. Group 1 (or A in the graph), group 2 (or B in the graph), Group 3 (or C in the graph) and Group 4 (or D in the graph) were assigned respectively to Tommy Atkins ripened on the tree and in storage and Haden ripened on the tree and in storage respectively. For the individual variables (the peak areas), three samples were misclassified from group 1 and one sample in group 3 resulting in 14% error of classification. Using the principal components, two samples were misclassified in group 1, three in group 3 and 1 in group 4, yielding an error of classification of 21%. Therefore, for the predictive equations and graphical representation of the
Table 4. Actual group membership (rows) vs. predicted (columns) for LDA of GC results using both individual and the first 3 principal components from PCA. Group 1 = Tommy Atkins ripened on the tree, Group 2 = Tommy Atkins ripened in storage, Group 3 = Haden ripened on the tree and Group 4 = Haden ripened in storage.
data (Figure 29) the individual component results were used. As in PCA, a 3-dimensional plot of the first three factors (or canonical variables) was used to try and classify the samples. In LDA, the groups were more distinguishable than in PCA and only two samples appear to be misclassified (sample A, Tommy Atkins lot 2 ripened on the tree individual 1, misclassified into group B and sample C, Haden lot 1 ripened on the tree individual 3, misclassified into group D). Although Table 4 shows 4 samples being incorrectly grouped, the other two samples have very similar probability of being in their rightful group or in the misclassified one.

The predictive equations (below) can be used to assign an unknown sample to a predetermined group given the seven peak areas (ie the sample belongs to the group whose equation yields the greatest positive value).

**Predictive Equations:**

GROUP 1 = 1.5A1 + 3.6A2 - 3.7A3 + 1.8A4 + 41.8A5 + 7.0A6 + 6.5A7 - 15.0
GROUP 2 = 2.4A1 - 3.4A2 - 0.4A3 + 0.2A4 + 55.3A5 + 28.4A6 - 6.5A7 - 10.4
GROUP 3 = 3.1A1 + 5.1A2 + 1.7A3 + 1.2A4 + 2.8A5 - 9.0A6 + 10.2A7 - 8.9
GROUP 4 = 9.5A1 + 8.1A2 - 1.7A3 + 12.4A4 - 21.4A5 - 29.3A6 + 10.2A7 - 9.8
where A1-A7 are the peak areas

3. **Principal Component Analysis - Similarity**

The first three factors from PCA and the % variance they each explain (Table 3b) were subjected to PCA-SIM. The arbitrary magnification factor (M) was chosen as 10.
Figure 29. Plot of factor scores (Canonical variables) resulting from LDA for the individual peak areas in the GC analysis. Symbols represent the following: A= Tommy Atkins ripened on the tree, B= Tommy Atkins ripened in storage, C= Haden ripened on the tree and D= Haden ripened in storage.
A Tommy Atkins sample (individual #4) of size 12 (lot 1) ripened in storage was chosen as a reference. This sample was selected at random since no ideal conditions were required for the comparison. The coefficients of determination as well as the slopes between the samples and the standard were calculated. The same four lots were used as in the previous classification methods.

Well defined groups for both varietal difference and method of ripening were formed and can be visualized in Figure 30. The reference (with a slope and coefficient of determination equal to one) is circled in the graph. Both varieties that ripened in storage formed tighter, more confined groups demonstrating less within group variation than those ripened on the tree. One Haden sample ripened on the tree was misclassified as a Haden ripened in storage and likewise one Tommy Atkins sample ripened on the tree was incorrectly grouped with the Tommy Atkins ripened in storage. In addition a Haden ripened in storage (lot 2, individual 3) did not belong to any of the groups. To account for the difference between this latter sample and the standard, the corrected PCA of the sample was plotted against the corrected PCA of the reference (Figure 31). The largest deviation was accounted for by the second principal component which is heavily loaded by peak areas one and three. Therefore, these peaks attribute to the substantial difference between the standard and this sample.
Figure 30. PCA-SIM results of the seven GC peak areas utilizing Tommy Atkins, lot 1, individual 4, ripened in storage as the standard (circled). The symbols represent the following: b= Tommy Atkins ripened in storage, c= Tommy Atkins ripened on the tree, o= Haden ripened in storage and k= Haden ripened on the tree.
Figure 31. Corrected Principal Component Scores (CPCS) of the standard (Tommy Atkins, lot 1, individual 4, ripened in storage) compared to the CPCS of the sample (Haden, lot 2, individual 3, ripened in storage). The 45 degree line would have signified a perfect match between the CPCS of the standard and CPCS of the sample.
Of the three classification methods, LDA and PCA-SIM were more successful in visually classifying the samples into the four groups using the gas chromatography results.

E. SENSORY ANALYSIS RESULTS

1. Taste

In the Magnitude Estimation trials, all judges ranked the sucrose solutions in the correct order and only one judge failed to rank the citric acid solutions correctly, therefore all judges were kept for the rest of the training.

The sweet and sour normalized and averaged (discussed below) sensory results for the two varieties are shown in Table 5 and Figure 32. Except for sample 4 for Tommy Atkins and sample 16 for the Haden, the samples ripened on the tree (terms ending in mm2 in Table 5) were rated sweeter than those ripened in storage (terms ending in g2 in Table 5) while sourness had the reverse trend. Comparison of the sensory and chemical results will follow the sensory section.

2. Aroma

Seven terms were decided upon by the panel members as adequate descriptors for the mango aroma and these were the following: overall intensity, tea-like, weed-like, fruity, pineapple or banana-like, honeydew melon-like and sweet. After several training and practice sessions the panellists were in agreement as to the meaning of each term and its intensity in several test samples.
<table>
<thead>
<tr>
<th>sample</th>
<th>Sample #</th>
<th>SweetT</th>
<th>Sour</th>
<th>S/A</th>
<th>MT(Kg)</th>
<th>pH</th>
<th>TA</th>
<th>Glu</th>
<th>Fru</th>
<th>Suc</th>
<th>TS</th>
</tr>
</thead>
<tbody>
<tr>
<td>t11gm2</td>
<td>1</td>
<td>2.72</td>
<td>8.46</td>
<td>29.88</td>
<td>1.40</td>
<td>3.83</td>
<td>0.34</td>
<td>0.41</td>
<td>2.16</td>
<td>7.59</td>
<td>10.16</td>
</tr>
<tr>
<td>t12gm2</td>
<td>2</td>
<td>4.99</td>
<td>5.88</td>
<td>42.42</td>
<td>1.22</td>
<td>4.17</td>
<td>0.26</td>
<td>0.44</td>
<td>2.36</td>
<td>8.23</td>
<td>11.03</td>
</tr>
<tr>
<td>t13gm2</td>
<td>3</td>
<td>4.71</td>
<td>5.76</td>
<td>31.00</td>
<td>1.26</td>
<td>4.09</td>
<td>0.34</td>
<td>0.40</td>
<td>2.51</td>
<td>7.63</td>
<td>10.54</td>
</tr>
<tr>
<td>t14gm2</td>
<td>4</td>
<td>7.97</td>
<td>0.74</td>
<td>----</td>
<td>1.09</td>
<td>3.78</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t11mm2</td>
<td>5</td>
<td>7.64</td>
<td>2.02</td>
<td>39.36</td>
<td>1.41</td>
<td>4.17</td>
<td>0.25</td>
<td>0.52</td>
<td>2.53</td>
<td>6.79</td>
<td>9.04</td>
</tr>
<tr>
<td>t12mm2</td>
<td>6</td>
<td>9.09</td>
<td>1.54</td>
<td>10.61</td>
<td>1.44</td>
<td>3.76</td>
<td>0.84</td>
<td>0.65</td>
<td>2.41</td>
<td>5.85</td>
<td>8.91</td>
</tr>
<tr>
<td>t13mm2</td>
<td>7</td>
<td>8.00</td>
<td>1.74</td>
<td>9.25</td>
<td>1.42</td>
<td>3.76</td>
<td>0.81</td>
<td>0.33</td>
<td>2.34</td>
<td>4.82</td>
<td>7.49</td>
</tr>
<tr>
<td>t14mm2</td>
<td>8</td>
<td>8.00</td>
<td>2.38</td>
<td>11.04</td>
<td>0.91</td>
<td>3.79</td>
<td>0.82</td>
<td>1.39</td>
<td>3.10</td>
<td>4.48</td>
<td>9.05</td>
</tr>
<tr>
<td>h11gm2</td>
<td>9</td>
<td>2.89</td>
<td>8.72</td>
<td>22.68</td>
<td>0.88</td>
<td>3.92</td>
<td>0.59</td>
<td>0.74</td>
<td>2.87</td>
<td>9.77</td>
<td>13.38</td>
</tr>
<tr>
<td>h12gm2</td>
<td>10</td>
<td>2.46</td>
<td>8.38</td>
<td>46.71</td>
<td>0.82</td>
<td>4.10</td>
<td>0.28</td>
<td>0.60</td>
<td>2.62</td>
<td>9.86</td>
<td>13.08</td>
</tr>
<tr>
<td>h13gm2</td>
<td>11</td>
<td>1.95</td>
<td>7.97</td>
<td>15.64</td>
<td>0.88</td>
<td>3.52</td>
<td>0.81</td>
<td>0.60</td>
<td>2.57</td>
<td>9.50</td>
<td>12.67</td>
</tr>
<tr>
<td>h14gm2</td>
<td>12</td>
<td>2.45</td>
<td>6.34</td>
<td>----</td>
<td>1.14</td>
<td>4.08</td>
<td>0.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h11mm2</td>
<td>13</td>
<td>6.66</td>
<td>2.22</td>
<td>11.05</td>
<td>0.98</td>
<td>3.96</td>
<td>0.81</td>
<td>0.51</td>
<td>2.01</td>
<td>6.43</td>
<td>8.95</td>
</tr>
<tr>
<td>h12mm2</td>
<td>14</td>
<td>7.26</td>
<td>2.32</td>
<td>15.06</td>
<td>0.82</td>
<td>3.72</td>
<td>0.62</td>
<td>0.05</td>
<td>2.11</td>
<td>7.18</td>
<td>9.34</td>
</tr>
<tr>
<td>h13mm2</td>
<td>15</td>
<td>7.32</td>
<td>2.56</td>
<td>20.50</td>
<td>0.79</td>
<td>3.72</td>
<td>0.54</td>
<td>0.69</td>
<td>2.62</td>
<td>7.76</td>
<td>11.07</td>
</tr>
<tr>
<td>h14mm2</td>
<td>16</td>
<td>4.14</td>
<td>5.63</td>
<td>67.39</td>
<td>0.66</td>
<td>4.52</td>
<td>0.18</td>
<td>0.64</td>
<td>2.81</td>
<td>8.68</td>
<td>12.13</td>
</tr>
</tbody>
</table>

Table 5. Averaged and normalized sensory results for 16 mango samples. Sample codes are explained in Appendix 2. The following abbreviations were used: SweetT= sweet taste, S/A= sugar-acid ratio, MT= Magness Taylor values, TA= titratable acidity, Glu= glucose, Fru= fructose, Suc= sucrose and TS= total sugars.
Figure 32. Averaged and normalized panel ratings for sweetness and sourness using QDA for Tommy Atkins and Haden mangos. Sample numbers were obtained from Table 5.
Normalization of raw data

For multivariate analysis, a representative set of results was needed, however, the raw data could not be simply averaged because of the large spread in the panelists' responses. Therefore, a normalization procedure suggested by Weiss and Zenz, in 1989, was applied to the raw data. This method involves two steps: in the first, the individual sensitivities of the judges were normalized to the averaged sensitivity of the panel and in the second, the zero point was normalized. The equations used were the following:

**Step 1:** \[ R_{i,\text{sens}} = \frac{\sum (R_{i,\text{max}}-R_{i,\text{min}})}{n (R_{i,\text{max}}-R_{i,\text{min}})} \]

**Step 2:** \[ R_{i,\text{norm}} = \frac{[\sum R_{i,\text{sens}} - (\sum R_{i,\text{sens}}) n]}{r} \]

where:

- \( R_{i,\text{sens}} \) = sensitivity normalized responses of the subjects
- \( i \) = subject ranging from 1 to \( n \)
- \( n \) = number of subjects in the panel
- \( R_{i,\text{max}}-R_{i,\text{min}} \) = range of the responses
- \( R_{i,\text{norm}} \) = normalized sensory response
- \( r \) = number of samples rated by each subject

Box plots of the raw and normalized data (Figures 33-41) demonstrate the effect of the normalization procedure. The Box plot graphically displays the distribution symmetry and central tendency of the data (Ott, 1988). Figure 42 explains the structure and symbols of the Box plot and Appendix 3 shows
Figure 33. Box plots of sensory data before and after normalization procedure for the parameter fruity. Treatment numbers were obtained from Table 5.
Figure 34. Box plots of sensory data before and after normalization procedure for the parameter weed. Treatment numbers were obtained from Table 5.
Figure 35. Box plots of sensory data before and after normalization procedure for the parameter sweet aroma. Treatment numbers were obtained from Table 5.
Figure 36. Box plots of sensory data before and after normalization procedure for the parameter pineapple/banana. Treatment numbers were obtained from Table 5.
Figure 37. Box plots of sensory data before and after normalization procedure for the parameter honeydew melon. Treatment numbers were obtained from Table 5.
Figure 38. Box plots of sensory data before and after normalization procedure for the parameter overall intensity. Treatment numbers were obtained from Table 5.
Figure 39. Box plots of sensory data before and after normalization procedure for the parameter sour taste. Treatment numbers were obtained from Table 5.
Figure 40. Box plots of sensory data before and after normalization procedure for the parameter tea. Treatment numbers were obtained from Table 5.
Figure 41. Box plots of sensory data before and after normalization procedure for the parameter sweet taste. Treatment numbers were obtained from Table 5.
Definitions: (Velleman and Haglin, 1981)

**Median** = Splits the ordered batch of numbers in half.

**Hinges** = Split the remaining halves in half again.

**H- Spreads** = The absolute value of the difference between the values of the hinges.

**Whiskers or Inner Fences** = Show the range of values within 1.5(Hspreads) of the hinges.

= lower hinge - 1.5(Hspreads)
= upper hinge - 1.5(Hspreads)

**Outer Fences** = lower hinge - 3(Hspreads)
= upper hinge - 3(Hspreads)

Any value beyond the whiskers are termed OUTSIDE and any data beyond either whiskers or outer fences are called FAR OUTSIDE.

Figure 42. Explanation of the Box plot structure and symbols.

102
a sample calculation using the results of treatment 1 for intensity (Tommy Atkins, lot 1, ripened on the tree, mix 2). Normalization narrowed the spread of the data and, therefore, the average of each treatment was representative of that treatment.

Table 6 displays the averages and standard deviations of the normalized data. For both Tommy Atkins and Haden varieties the attributes of overall intensity, fruity, pineapple/banana, honeydew, sweet aroma and taste were higher in the samples ripened on the tree (those ending in MM2) than in the ones ripened in storage (those ending in G2), while tea, weed and sourness were ranked higher in the mangos ripened in storage rather than those ripened on the tree. The only exceptions were lot 4 for Tommy Atkins and lot 4 for Haden. The more "green-like" characteristics of weed, tea-like and sourness were, therefore, more characteristic of the fruit ripened in storage. Multivariate analysis was subsequently applied to establish if these differences would be sufficient to classify the samples into their appropriate groups.

F. MULTIVARIATE ANALYSIS OF SENSORY RESULTS

Principal Component Analysis

PCA was performed on the averaged, normalized taste and aroma results (nine variables in all). Only the first eigenvalue gave a value greater than one (7.61) and the first
Table 6. Averages and standard deviations of the normalized results for the nine sensory attributes. The treatment encoding may be referred to in Appendix 2.
principal component (factor 1) explained 84.5% of the total variance. Table 7 shows the component loadings of each variable for factor 1 and factor 2. The first principal component scores were plotted for each group (groups 1, 2, 3 and 4 were respectively, Tommy Atkins ripened in storage and on the tree, and Haden ripened in storage and on the tree) shown in Figure 43. Groups that appeared in the positive side of the graph were ripened on the tree and those on the negative side of the graph were ripened in storage.

Groups 2 and 3 were well defined while two samples from group 1 and one sample from group 4 were misclassified. The samples from group 1 (lot 1 incorrectly classified into group 3 (purple) and lot 4, misclassified into group 2 (green)) and from group 4 (lot 4 incorrectly classified into group 3 (purple)) had similar values for the descriptive parameters as the group they were misclassified into (Table 6). PCA with 20% error in classification was not considered very successful in differentiating variety and ripening conditions given the sensory results.

2. Linear Discriminant Analysis

As for the gas chromatography results, LDA was performed both on the nine individual sensory descriptive terms as well as the two first principal components obtained from PCA of the sensory terms. Groups 1-4 were the same as used in the PCA analysis. Table 8 shows that only one sample from group 3 was
<table>
<thead>
<tr>
<th>Descriptive Variable</th>
<th>Component Loadings</th>
<th>Eigenvalue</th>
<th>% Total Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
<td>Factor 2</td>
<td></td>
</tr>
<tr>
<td>SweetA</td>
<td>0.979</td>
<td>0.018</td>
<td>7.61</td>
</tr>
<tr>
<td>Fruity</td>
<td>0.978</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>PinBan</td>
<td>0.970</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>SweetT</td>
<td>0.967</td>
<td>-0.035</td>
<td></td>
</tr>
<tr>
<td>Sour</td>
<td>-0.941</td>
<td>0.058</td>
<td></td>
</tr>
<tr>
<td>Weed</td>
<td>-0.920</td>
<td>0.186</td>
<td></td>
</tr>
<tr>
<td>Tea</td>
<td>-0.901</td>
<td>0.186</td>
<td></td>
</tr>
<tr>
<td>Honeydew</td>
<td>0.826</td>
<td>-0.150</td>
<td></td>
</tr>
<tr>
<td>Intensity</td>
<td>0.767</td>
<td>0.633</td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Sensory results of PCA
Figure 43. Sensory PCA results of Principal Component 1. Group 1= Tommy Atkins ripened in storage, Group 2= Tommy Atkins ripened on the tree, Group 3= Haden ripened in storage and Group 4= Haden ripened on the tree. Colors designate the group into which the sample was classified.
misclassified (6% error of classification) using the individual variables for LDA while two samples in group 1 and two samples in group 4 were incorrectly classified (25% error of classification) using the principal components for LDA. Therefore, for the predictive equations and graphical results the individual variables data was used. Distinct groups were formed when the three canonical variables (factors 1-3) were plotted against each other (Figure 44). Variables A through D correspond respectively to groups 1 through 4 in the PCA analysis. The LDA results indicate that the sensory panel was, on average, able to distinguish between the two varieties and the two ripening conditions.

The predictive equations for the descriptive terms were the following:

GROUP 1 = \[ 85A + 262B - 234C + 87D + 303E - 82F - 43G + 169H + 289I - 1848 \]

GROUP 2 = \[ 125A + 334B - 318C + 129D + 426E - 121F - 94G + 207H + 373I - 3010 \]

GROUP 3 = \[ 89A + 271B - 243C + 92D + 315E - 81F - 49G + 169H + 297I - 1945 \]

GROUP 4 = \[ 89A + 270B - 242C + 94D + 312E - 79F - 56G + 172H + 293I - 1916 \]

where:

- A = intensity,
- B = tea-like,
- C = weed-like,
- D = fruity,
- E = pineapple/banana,
- F = honeydew melon-like,
- G = sweet aroma,
- H = sweet taste and
- I = sour taste

3. **Principal Component Analysis - Similarity**

PCA-SIM was not performed on the sensory data since only the first principal component was needed to explain the variance in the data.

108
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 8. Actual group membership (rows) vs. predicted (columns) for LDA of the individual descriptive terms as well as principal components from PCA. Group 1 = Tommy Atkins ripened in storage, Group 2 = Tommy Atkins ripened on the tree, Group 3 = Haden ripened in storage and Group 4 = Haden ripened on the tree.
Figure 44. Plot of factor scores (canonical variables) resulting from LDA for the individual sensory descriptive terms. Symbols represent the following: A= Tommy Atkins ripened in storage, B= Tommy Atkins ripened on the tree, C= Haden ripened in storage and D= Haden ripened on the tree.
One of the main functions of PCA-SIM is to simplify the visualization of the data when there are three or more variables (principal components) involved. Therefore, PCA-SIM was not applicable in this case.

Of the three multivariate analysis techniques, LDA was most appropriate to classify the mangos based on their varietal and ripening conditions utilizing the sensory results.

G. COMPARISON OF CHEMICAL AND SENSORY DATA FOR TASTE

For the 16 samples compared, the titratable acidity ranged from 0.18% to 0.84% citric acid while total sugar content ranged from 7.49% to 13.38% (Table 5). Pairi mangos contained 0.16% citric acid after 21 days of storage at 15 C (Abou Aziz et al., 1975). Mangos harvested mature but green and ripened at room temperature (25-27 C) contained 0.18% and 0.13% citric acid for Tommy Atkins and Haden varieties (Peacock and Brown, 1986) and a range from 0.11 to 1.33% (Lodh et al., 1974), 0.18 to 0.45% (Selvaraj and Pal, 1988) and 0.10 to 0.62% (Kulkarni and Rameshwar, 1981) citric acid for several Indian varieties. Total sugars were found to be 4-6% for Baneshan mangos using the Lane and Eynon methods (Suryaprakasa Rao et al., 1972), 7.35- 13.20% using the Phenol Sulfuric Acid method (Lodh et al., 1974) and 9.3- 21% using
Fehling reagent method (Kulkarni and Rameshwar, 1981) for several Indian varieties, 13.1% for Pairi mangos stored at 26 C for 14 days using Somogyi’s method (Krishnamurthy and Subramanyam, 1970) and 27.6% for Pairi mangos stored at 15 C for 20 days. Some of the discrepancy between these literature values and the current study may be attributed to the method used for sugar determination, varietal differences as well as storage temperatures.

Sucrose was found to be the predominant sugar followed by fructose and glucose. Sharaf and co-workers in 1989 found that at the green mature stage, sucrose is most abundant followed by fructose and glucose while in the ripe and overripe stages glucose and fructose respectively, predominated. However, Tandon and colleagues in 1973 found that for Dashehari mangos glucose and fructose were the main sugars in the early fruit development and changed to glucose and sucrose toward ripening. The differences between the results obtained in this work and that of other authors may be the same as those discussed for the total sugar determination.

The sugar-acid ratios ranged from 9.25 to 67.39 with no apparent trend. The ratio was 83.2 and 1.34 for Tommy Atkins and Haden respectively (Peacock et al., 1986) and ranged from 5.5 to 109.2 (Lodh et al., 1974) and 30.48 to 71.01 (Selvaraj and Pal, 1988) for several Indian varieties. No correlation could be drawn between the ratio and the sensory results. Therefore, the sugar-acid ratio was unable to distinguish
between varietal differences nor between ripening conditions in the present study.

H. COMPARISON OF SENSORY TO GC DATA

No correlation (a probability less than 95%) was found between the seven individual peak areas or the three principal components of the peak areas and the sensory descriptive terms. No general tendency was apparent for any of the descriptive terms and the GC data (Figures 45 and 46). Martens and co-workers were also unable to find a relationship between sensory and chemical data for their product (raw carrots), and Powers in 1981 discussed at length the difficulties in attempting such correlations. However, conclusive results cannot be drawn since only one sample from each lot was tested.

I. PH AND MAGNESS TAYLOR RESULTS

Figure 47 shows the pH and Magness Taylor results for the samples used in the sensory analysis. Neither parameter demonstrated a significant trend in these samples. The pH of the samples fluctuated in the acidic range between 3.52 and 4.52 and the Magness Taylor values ranged from 0.66 to 1.44Kg (Table 5). Medlicott and co-workers in 1990 found the Magness Taylor values for Tommy Atkins mangos harvested mature but green and stored until ripe (for 21 days) at 12 C to range between 0.91 to 1.56Kg which are comparable to the values
Figure 45. Scattergrams of individual peak areas (PA1-PA7) from GC analysis with individual sensory descriptive terms: an attempt to find a correlation between sensory and GC data.
Figure 46. Scattergram of factor scores from PCA of GC results with individual sensory descriptive terms: an attempt to find a correlation between sensory and GC data.
Figure 47. Magness Taylor values and pH results of the samples used for sensory analysis. Sample numbers were obtained from Table 5.
obtained in this study. For pH, Tommy Atkins and Haden picked green and mature and ripened at 25 C and 80-90% RH had a value of 4.91 and 5.05, respectively (Peacock and Brown, 1986). Similarly, Haden mangos ripened at 27 C had a pH value of 4.3 after 5 days in storage (Soule and Harding, 1956) and African mangos had a pH ranging from 4.5 to 5.0 after 7 days in storage (Aina, 1990). The slightly higher values for the pH obtained in these studies may be due to the ripening temperature being much higher than in the present work.

The pH and Magness Taylor values for the averages per lot for all the samples are presented in Table 9. The Mangness Taylor value is higher (requiring more force= harder fruit) for the Tommy Atkins mangos ripened in storage (ending in G) when compared within lots to those ripened on the tree (ending in M). The only exception is lot 4. No such trend was seen in the Haden mangos, although overall, they had lower MT values than the Tommy Atkins.

The pH values were lower for the mangos ripened in storage than those ripened on the tree in both varieties. The only exception was lot 4 for Haden which interestingly enough had a very low MT value for the mangos ripened in storage. Therefore, given the averages for each lot, a possibility exists for classifying varietal differences with Magness Taylor values and ripening conditions with the pH. However, when these two variables were combined with sensory and GC data and subjected to PCA and LDA, no classification was
<table>
<thead>
<tr>
<th>Sample</th>
<th>MT</th>
<th>PH</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAL1M</td>
<td>1.20</td>
<td>4.28</td>
</tr>
<tr>
<td>TAL1G</td>
<td>1.44</td>
<td>3.74</td>
</tr>
<tr>
<td>TAL2M</td>
<td>1.08</td>
<td>4.45</td>
</tr>
<tr>
<td>TAL2G</td>
<td>1.39</td>
<td>3.70</td>
</tr>
<tr>
<td>TAL3M</td>
<td>1.14</td>
<td>4.24</td>
</tr>
<tr>
<td>TAL3G</td>
<td>1.64</td>
<td>4.13</td>
</tr>
<tr>
<td>TAL4M</td>
<td>1.06</td>
<td>3.96</td>
</tr>
<tr>
<td>TAL4G</td>
<td>0.99</td>
<td>3.74</td>
</tr>
<tr>
<td>HL1M</td>
<td>1.07</td>
<td>4.21</td>
</tr>
<tr>
<td>HL1G</td>
<td>0.90</td>
<td>3.79</td>
</tr>
<tr>
<td>HL2M</td>
<td>0.74</td>
<td>4.24</td>
</tr>
<tr>
<td>HL2G</td>
<td>0.82</td>
<td>3.86</td>
</tr>
<tr>
<td>HL3M</td>
<td>0.87</td>
<td>4.06</td>
</tr>
<tr>
<td>HL3G</td>
<td>0.87</td>
<td>3.74</td>
</tr>
<tr>
<td>HL4M</td>
<td>1.14</td>
<td>4.23</td>
</tr>
<tr>
<td>HL4G</td>
<td>0.70</td>
<td>4.63</td>
</tr>
</tbody>
</table>

Table 9. Averages for each lot of pH and Magness Taylor values. Coding for each sample was the following: TA= Tommy Atkins, H= Haden, L1 to L4= lot numbers, M= ripened on the tree and G= ripened in storage.
apparent. Suryaprakasa Rao and co-workers in 1970 while studying the maturity of Baneshan mangos, also could not find a trend using the Mangness Taylor pressure tester.

J. WEIGHT LOSS IN MANGOS

The average weights and standard deviations for both varieties of mangos and their respective lots are shown in Table 10. The % weight loss was calculated for the mangos ripened in storage. Within each variety, the % weight loss was consistent: 5% for Haden and 3-4% in Tommy Atkins. The difference between weight loss does not seem substantial between varieties. This weight loss is attributed, for the most part to water loss during storage. No weight difference was calculated for the mangos ripened on the tree since no initial weight was available (mangos were still on the tree). The average weight decreased from lot 1 to lot 3 since size decreased. For Haden mangos harvested green but mature and ripened at 27 C, the % weight loss was 4 and 5% after 5 and 6 days respectively (Soule and Harding, 1956). For the Keith variety harvested green but mature and ripened for two weeks at 13 C, the weight loss was 5.91% (Fornaris-Rollan et al., 1990). Also for African mangos ripened in storage at 27-30 C and 68-70% RH, the weight loss was 4 and 5% after 4 and 5 days respectively. Because conditions were slightly different in these publications as compared to the present study, it was difficult to compare the results.
<table>
<thead>
<tr>
<th>Sample</th>
<th>Avg</th>
<th>Std</th>
<th>% Wt. loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAL1b</td>
<td>440</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>TAL1a</td>
<td>423</td>
<td>27</td>
<td>4</td>
</tr>
<tr>
<td>TAL2b</td>
<td>365</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>TAL2a</td>
<td>352</td>
<td>30</td>
<td>4</td>
</tr>
<tr>
<td>TAL3b</td>
<td>307</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>TAL3a</td>
<td>298</td>
<td>24</td>
<td>3</td>
</tr>
<tr>
<td>TAL1</td>
<td>442</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>TAL2</td>
<td>367</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>TAL3</td>
<td>309</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>TAL4</td>
<td>340</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>HL1b</td>
<td>389</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>HL1a</td>
<td>370</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>HL2b</td>
<td>360</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>HL2a</td>
<td>342</td>
<td>23</td>
<td>5</td>
</tr>
<tr>
<td>HL3b</td>
<td>313</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>HL3a</td>
<td>298</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>HL1</td>
<td>435</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>HL2</td>
<td>383</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>HL3</td>
<td>327</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>HL4</td>
<td>307</td>
<td>39</td>
<td></td>
</tr>
</tbody>
</table>

Table 10. Average weights and standard deviations of mangos before (b) and after (a) storage (for those ripened in storage) and after picking (for those ripened on the tree). The % weight loss was calculated for mangos ripened in storage. Sample codes are explained in Appendix 2.
V. CONCLUSION

The demand for mangos has increased outside of mango-growing areas resulting in export, mostly by ground transportation, of this delicate fruit. The goal of the mango grower is to obtain mangos that will ripen evenly, and develop the proper flavour and textural properties (as compared with tree ripened fruit) during transport. Because aroma is such an important part of fruit flavour, it seemed a reasonable parameter for comparing tree ripened fruit to that ripened in cold storage. Gas chromatography and sensory analysis were used in this study to compare these conditions. In addition, physico/chemical parameters such as pH, sugar-acid ratio, weight loss and Magness Taylor pressure values were compared.

In order to show the differences between varieties (Haden and Tommy Atkins) and ripening conditions (ripened in storage and on the tree), three multivariate analysis techniques were used for the data analysis of the sensory and GC results. The first two methods, Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are common classification techniques in food science. The third method, Principal Component Analysis-Similarity (PCA-SIM) was developed in our laboratory and used as a comparison to the other techniques.

A new portable gas chromatograph (SRI Model 8610) equipped with a FID detector was used for the aroma analysis of the mango samples. The preliminary work resulted in choosing the method of analysis (purge and trap), the column
(DB-624), the trap (Tenax) and the temperature program (begin at 35 °C and hold for 10 min than ramp at 4 deg/min until 210 °C and hold for 5 min). Random Centroid Optimization was utilized for the optimization of the purge conditions, and after 18 experiments using Tommy Atkins mangos, best results were obtained when a 0.11 gram sample was heated at 51 °C for 8 min. These conditions were, therefore, used for the rest of the GC analysis.

After all the samples of Tommy Atkins and Haden varieties, individual and mixed, of different sizes and ripening conditions were analyzed by GC, seven major peaks were chosen in an attempt to classify the two varieties and two ripening conditions using the three multivariate analysis techniques. PCA was more successful in variable reduction than classification while LDA, using the individual variables, and PCA-SIM showed distinct differences between the groups.

Mass spectroscopy was used to identify the primary GC peaks. Car-3-ene, α-Pinene and β-Phellandrene were found to be the major compounds present in the headspace of these mangos.

Sensory analysis utilizing Quantitative Descriptive Analysis was also used to distinguish between varieties and ripening conditions. Panel members arrived at the following seven terms to describe the mango aroma: overall intensity, weed-like, tea-like, fruity, pineapple/banana like, honeydew melon like and sweet. The results of these seven aroma
parameters were combined with the sweet and sour taste results and subjected to the three multivariate analysis programs to attempt classification. PCA was unsuccessful in classifying the samples while LDA utilizing the individual descriptive terms formed four distinct groups. PCA-SIM was not used for the sensory data since one principal component was sufficient in explaining the majority of the variance.

From the three multivariate analysis techniques applied in this work, PCA was best used for variable reduction while LDA and PCA-SIM were successful at classifying the two mango varieties ripened in two different conditions. The PCA-SIM program was found to have the following advantages as compared to the currently available multivariate analysis techniques. Firstly, the PCA-SIM method is the only technique that allows the classification of samples in relation to a preselected standard. Secondly, regardless of the number of variables involved, the results can always be visualized in two dimensions (S vs. R plots). Finally, the concept of slope and coefficient of determination are commonly used and, therefore, easy to comprehend.

For both varieties, the mangos ripened in storage had a higher sourness and lower sweetness than those ripened on the tree. However, no correlation could be drawn between the chemical analysis and sensory qualities for sweet and sour.

Both techniques were successful in classifying the samples by varietal differences and ripening condition so
either may be used for this purpose. Individual GC peaks and PCA factor scores generated from these peaks could not be correlated to sensory attributes. However, in many cases more than one volatile contributes to a sensory perception. The correct combination of volatiles (if it exists) for each sensory attribute is needed in order for GC results to replace sensory analysis.

For the physico/chemical properties investigated, the size, weight and sugar-acid ratio of the fruit did not show any significant trends between variety or ripening condition. The pH differed between those mangos ripened on tree and those ripened in storage and the Magness Taylor value showed differences between varieties but neither parameter could distinguish between all four conditions.

RECOMMENDATION: J & W Scientific has recently made available a 75 meter option of the DB624 megabore column. Since the SRI gas chromatograph cannot be used with cryo-focusing, only with cryo-trapping, better resolution may be obtained with this longer column. When tested in our laboratory, the longer megabore column showed an improved resolution and number of peaks when compared to the 30 meter capillary DB624 column. The better resolution was not only due to the length per se (i.e. resolution being a function of the square root of the column length) but also the prevention of peak broadening due to higher head pressure. The megabore column also allows for
higher flow rates which are important in the desorption and subsequent transfer of the volatiles to the column. Further work is required (including re-optimization of all the GC parameters) to verify the usefulness of this new column.
VI. REFERENCES


Appendix 1. Principal Component Analysis - Similarity computer program.

10 DIM M(30), L(90)
20 DIM N(90), CC(90)
30 DIM A(90,50), R(90)
40 DIM B(20,90)
50 DIM REF(10), S(50)
60 DIM VARB(30), K(30)
70 DIM Q(90,20), Y(90)
80 DIM X(90), PDX(40), POY(40)
90 DIM X1(50), Y1(50)
92 CLS: SCALES$ = "CDEFGAB": PLAY "04 XSCALES;"
95 SCREEN 1: COLOR 12: KEY OFF
100 LOCATE 10, 40: PRINT "^ PCA-SIMILARITY ANALYSIS": PRINT : PRINT
110 INPUT "^ PRESS ANY KEY TO CONTINUE", ZZ: CLS
115 SCREEN 2: SCREEN 0
130 INPUT "FILE NAME TO BE USED"; FILE$
140 INPUT "HOW MANY CASES IN EACH FACTOR"; C
150 INPUT "HOW MANY FACTORS WOULD YOU LIKE TO USE"; F
160 INPUT "WHICH CASE (ROW NUMBER) DO YOU WANT TO BE YOUR STANDARD"; S
170 INPUT "WHAT WOULD YOU LIKE YOUR ARBITRARY MAGNIFICATION TO BE"; M
180 FOR X = 1 TO F
190 PRINT "WHAT IS THE % VARIANCE FOR FACTOR NUMBER"; X
200 INPUT VARB(X)
210 NEXT
220 REM INPUTING FILES FROM PCA
230 OPEN FILE$ FOR INPUT AS #1
240 FOR I = 1 TO C
250 FOR J = 1 TO F
260 INPUT #1, A(I,J)
270 NEXT J
280 NEXT I
290 CLOSE #1
295 REM REFERENCE USED IN CALCULATION
300 FOR J = 1 TO F
310 REF(J) = A(S, J)
320 NEXT J
330 REM conversion to 100 Dr. Nakai's idea
334 FOR I = 1 TO F
338 RVARB(I) = VARB(I)
339 NEXT I
340 REM SIMILARITY ANALYSIS EQUATION
350 FOR I = 1 TO C
360 FOR J = 1 TO F
370 IF J = 1 THEN S(J) = 100
380 K(J) = S(J) - RVARB(J)
390 S(J+1) = K(J)
400 IF J = 1 THEN Q(I,J) = ((A(I,J) - REF(J)) * M) + 100: PRINT A(I,J), REF(J), Q(I,J): INPUT LL
410 IF J = 1 THEN 430
420 Q(I,J) = ((A(I,J) - REF(J)) * M) + S(J): PRINT A(I,J), REF(J), S(J), Q(I,J): INPUT "ANOTHER"; LL
430 NEXT J
440 NEXT I
450 FOR I = 1 TO F
460 FOR J = 1 TO C
470 B(I,J) = Q(J, I)
480 NEXT J
490 NEXT I
500 REM SUBROUTINE: REGRESSION
510 FOR P = 1 TO C
520 SX = 0: SY = 0: SX2 = 0: SY2 = 0: SXY = 0
530 FOR I = 1 TO F
540 Y(I) = B(I, P)
550 X(I) = B(I, S)
560 NEXT I
570 FOR T = 1 TO F
580 SX = SX + X(T)
590 SY = SY + Y(T)
600 SX2 = SX2 + X(T)^2
610 SY2 = SY2 + Y(T)^2
620 SXY = SXY + X(T) * Y(T)
630 NEXT T
640 REM SLOPE
650 L(P) = (C * SXY - SX * SY) / (C * SX2 - SX^2)
660 REM COEFFICIENT OF CORRELATION
670 CC(P) = (SXY - SX * SY / C) / (SQR((SX2 - (SX^2) / C) * (SY2 - (SY^2) / C)))
680 R(P) = CC(P) ^ 2
690 NEXT P
700 LPRINT "CASE", "SLOPE", "COEF. OF DETERM."
710 REM GRAPHICS
720 FOR P = 1 TO C
730 X(P) = R(P)
740 Y(P) = L(P)
750 LPRINT P, L(P), R(P)
760 PRINT P, L(P), R(P): NEXT P
770 INPUT "NAME OF FILE FOR R^2 AND SLOPE"; FILS$
780 OPEN FILS$ FOR OUTPUT AS #1
790 FOR I = 1 TO C
800 PRINT #1, X(I), Y(I)
810 NEXT I
820 XMIN = 0
830 YMIN = 0
840 XMAX = 1
850 YMAX = 3
860 NPTS = C
870 XLABS$ = "COEFFICIENT OF DETERMINATION"
880 YLAB$ = "SLOPE"
890 SYM = 1
900 SIZE = .5
910 XLEN = 0
920 YLEN = 0
930 XLIN = 0
940 YLIN = 0
950 XINC = 0
960 MORE = 0
970 GOSUB 1430
980 LOCATE 1, 2: INPUT "PRESS ANY KEY TO CONTINUE (PRESS PRINT SCREEN TO PLOT GRAPH)"; ZZ:CLS
990 A$ = "N"
1000 IF A$ = "N" THEN 1080
1010 IF A$ = "Y" THEN 1020
1020 PRINT "THESE ARE THE CURRENT X MIN. AND X MAX."; XMIN, XMAX
1030 INPUT "ENTER THE NEW X MIN, X MAX"; XMIN, XMAX
1040 PRINT "THESE ARE THE CURRENT Y MIN. AND Y MAX."; YMIN, YMAX
1050 INPUT "ENTER THE NEW Y MIN, Y MAX"; YMIN, YMAX
1060 GOSUB 1430
1070 LOCATE 1, 2: INPUT "PRESS ANY KEY TO CONTINUE (PRESS PRINT SCREEN TO PLOT GRAPH)"; ZZ:CLS
135
1080 INPUT "WOULD YOU LIKE TO SEE THE GRAPH OF ANY SPECIFIC CASE (Y/N)"; Y$
1090 IF Y$ = "Y" THEN 1110
1100 IF Y$ = "N" THEN 1410
1110 INPUT "WHICH CASE # WOULD YOU LIKE TO SEE"; CN
1120 XMIN=0: XMAX=0: YMIN=0: YMAX=0
1125 PRINT : PRINT TAB(15) "STANDARD= "; S; : PRINT TAB(15) "VS. SAMPLE= "; CN
1130 FOR I = 1 TO F
1140 X(I) = B(I, S): PRINT: PRINT "CORRECTED PCA("; I; ")"; : PRINT TAB(17) B(I, S);
1150 Y(I) = B(I, CN): PRINT TAB(30) B(I, CN)
1160 IF X(I) > AMAX THEN AMAX = X(I)
1170 IF X(I) < AMIN THEN AMIN = X(I)
1180 IF Y(I) > BMAX THEN BMAX = Y(I)
1190 IF Y(I) < BMIN THEN BMIN = Y(I)
1200 NEXT I
1205 PRINT: INPUT "DO YOU WANT TO SAVE THE CORRECTED PCA"; MM$
1207 IF MM$ = "N" THEN 1210
1208 INPUT "FILE NAME TO SAVE RESULTS"; NOM$: OPEN NOM$ FOR OUTPUT AS #1: FOR I = 1 TO F: PRINT #1, B(I, CN), B(I, S): NEXT I: CLOSE
1210 XMAX = AMAX
1220 XMIN = AMIN
1230 YMAX = BMAX
1240 YMIN = BMIN
1250 NPTS = F
1260 LAB1$ = "CASE #"
1270 CORNER = 1
1280 XLAB$ = "STANDARD"
1290 YLAB$ = "CASE"
1300 LTYPE = 0
1310 MORE = 0
1320 GOSUB 1430
1330 MORE = 1: X(1) = XMAX
1340 X(2) = XMIN
1350 Y(1) = X(1)
1360 Y(2) = X(2)
1370 LTYPE = 1
1380 NPTS = 2
1390 GOSUB 1430
1400 GOTO 1070
1410 PRINT: PRINT: PRINT TAB(30) "--------END OF PROGRAM--------": END
1420 'SUBGRAPH
1430 "SUBGRAPH is a BASIC subroutine that will plot any set of data.
1440 'It defines the coordinates of the screen, draws
1450 'the grid with tic marks, numbers, and labels on each axis.
1460 'It also prints a 3 line label in one corner. The user can specify
1470 'the size and proportion of the figure in inches. Either linear
1480 'or logarithmic axes are allowed on both ordinate and abcissa.
1490 'Alt-C controls the color of the plot, the entire figure is
1500 'plotted in the chosen color. The color option does not
1510 'affect the printer dump print density.
1520 'Calling sequence:
1530 ' 1. MERGE "SUBGRAPH.bas" with your program
1540 ' 2. Save a new copy of your program which includes subroutine
1550 ' 3. Set up the parameters as defined below
1560 ' 4. gosub 60000 at point where you want the plot
1570 '---------------------------------------------------------------
'Parameters: Do not use these names elsewhere in your program

SYM symbol type (0=none,1=open sq,2=fill sq,3=open tri
4=fill tri,5=open cir,6=fill cir
7=open diamond,8=filled diamond,9=X

SIZE Symbol size in % of axes length

LTYPE line type (0=none,1=solid,2=dashed,3=dotted,4=regression
5=scatter)

NTS is the number of data points

X(I), Y(I) arrays that contain the x and y data points

XLEN,YLEN x and y axis length in inches

XMIN,YMIN x and y minimum values

XM, YM X and y axis maximum values

XLIN,YLIN flag 1 or linear (=0) or Log(=1) axis

XINC,YINC unit increment on each axis (valid only for linear)

MORE if these are <=0 they are calculated from data

LAB$, YLAB$ strings containing the axis labels

CORNER specifies the corner for the label

(0=none,1=LL,2=LR,3=UL,4=UR)

MORE a flag that indicates whether this call is the
first and so axes should be plotted, or if it
is more data to go on the same axes (axes are
not plotted if MORE=1), MORE=0 new figure.

'MIRORSAVFIL'

For more than one set of data on the same axes:

1. Set up the first set of data and all other plot parameters
2. GOSUB 60000 with MORE=0
3. Set up second set of data (leave plot parameters unchanged)
4. GOSUB 60000 with MORE=1
5. repeat steps 3 and 4 for each additional data set

Scale axes and plot them

IF YLEN<0 THEN YLEN=5.5
IF XLEN<0 THEN XLEN=7.5
IF MORE<1 THEN SCREEN 2:CLS:KEY OFF
KEY 20,CHR$(48)+CHR$(46):KEY (20) ON
ON KEY (20) GOSUB 3460
IF KLR.P=0 THEN KLR.P=15
OUT 985,KLR.P
XINC.P=XINC:YINC.P=YINC
XRANGE.P=XMAX-XMIN:YRANGE.P=YMAX-YMIN
IF XINC<0 THEN XINC.P=10^((INT(LOG(XRANGE.P*.66)/LOG(10))))
IF YINC<0 THEN YINC.P=10^((INT(LOG(YRANGE.P*.66)/LOG(10))))
YMIN.P=YINC.P*INT(YMIN.YINC.P+1):YMIN.P=YINC.P*INT(YMIN.YINC.P)
XMIN.P=YINC.P*INT(XMIN/XINC.P):XMAX.P=XINC.P*(INT((XMAX/XINC.P)+1))
YMAX.P=YINC.P*INT((YMAX/YINC.P)+1):YMIN.P=YINC.P*INT(YMIN/YINC.P)
IF XLIN=1 THEN XM=LOG(XMAX)/LOG(XMIN):YMIN=LOG(XMIN)/LOG(XMIN)
IF YLIN=1 THEN YM=LOG(YMAX)/LOG(YMIN):YM=LOG(YMIN)/LOG(YMIN)
XMIN.P=XMAX.P:YRANGE.P=YMAX.P-YMIN.P
DX=SIZE*XRANGE.P/100!:DY=SIZE*YRANGE.P/100!
X.T.P=XRANGE.P*(9!/YLEN):Y.T.P=YRANGE.P*(7!/YLEN)
TICK=.03*XRA.NE.P:TICY=.04*YRANGE.P
XTRA=(X.T.P-XRANGE.P)*9/X.T.P:YTRA=(Y.T.P-YRANGE.P)*7/Y.T.P
LBD.X=XMIN.P-(1!)*X.T.P/9
LBD.Y=YMIN.P-(1!)*Y.T.P/7
UBD.X=XMAX.P+((XTRA-1!)*X.T.P/9):UBD.Y=YMAX.P+((YTRA-1!)*Y.T.P/7)
2230 IF MORE<>1 THEN WINDOW (LBD.X,LBD.Y)-(UBD.X,UBD.Y)
2240 IF MORE<>1 THEN LINE (XMIN.P,YMIN.P)-(XMAX.P,YMAX.P),1,B
2250 XLOW.P=XMIN.P-LBD.X,YLOW.P=YMIN.P-LBD.Y
2260 XHI.P=XT.P-XRANGE.P-XLOW.P,YHI.P=YT.P-YRANGE.P-YLOW.P
2270 XP.P=.00159*(UBD.X-LBD.X)
2280 IF MORE<>1 THEN LINE (XMIN.P+XP.P,YMIN.P)-(XMAX.P+XP.P,YMAX.P),1,B
2290 STYLE=&HEFFF:IF LTYPE=0 THEN STYLE=&H0
2300 IF LTYPE=2 THEN STYLE=&HFOFO
2310 IF LTYPE=3 THEN STYLE=&HCOCO
2320 IF LTYPE=4 THEN STYLE=&H0
2330 IF MORE=1 THEN 3170
2340 'label axes
2350 XPOS.P=((XLOW.P+(XRANGE.P/2!))*80!)/XT.P)-(LEN(XLABS)/2)
2360 LOCATE 25,XPOS.P:PRINT XLAB$;
2370 YPOS.P=25!-(25!*((YLOW.P+(YRANGE.P/2!))/YT.P))-(LEN(YLABS)/2!)
2380 FOR I=1 TO LEN(YLABS):YT$=MID$(YLABS,I,1):LOCATE I+YPOS.P,3:PRINT YT$;:NEXT I
2390 'tic marks and numbers on linear x axis
2400 IF XLIN=1 THEN 2710
2410 MAXLEN=0:IF LEN(LAB3$)>MAXLEN THEN MAXLEN=LEN(LAB3$)
2420 IF LEN(LAB2$)>MAXLEN THEN MAXLEN=LEN(LAB2$)+1
2430 IF LEN(LAB1$)>MAXLEN THEN MAXLEN=LEN(LAB1$)+1
2440 IF CORNER=0 THEN GOTO 2590
2450 MAXLEN=0:IF LEN(LAB3$)>MAXLEN THEN MAXLEN=LEN(LAB3$)
2460 IF LEN(LAB2$)>MAXLEN THEN MAXLEN=LEN(LAB2$)+1
2470 IF LEN(LAB1$)>MAXLEN THEN MAXLEN=LEN(LAB1$)+1
2480 IF CORNER=1 OR CORNER=2 THEN XPOS.P=((XLOW.P/XT.P)*80!)+3
2490 IF CORNER=3 OR CORNER=4 THEN XPOS.P=((XLOW.P+XRANGE.P)/XT.P)*80!)-MAXLEN
2500 IF CORNER=2 OR CORNER=4 THEN XPOS.P=(((XLOW.P/XT.P)*80!)+3)
2510 IF CORNER=3 OR CORNER=4 THEN YPOS.P=((YHI.P/YT.P)*26)+2
2520 IF CORNER=1 OR CORNER=3 THEN YPOS.P=((YHI.P+YRANGE.P)/YT.P)*26!)-4!
2530 LOCATE YPOS.P+1,XPOS.P:PRINT USING "##.#";YTIC
2540 LOCATE YPOS.P+2,XPOS.P:PRINT USING "##.#";XTIC;
2550 NEXT YTIC
2560 'tic marks and numbers on linear y axis
2570 IF YLIN=1 THEN 2710
2580 FOR YTIC=YMIN.P TO YMAX.P STEP YINC.P
2590 LINE (XTIC,YMIN.P)-(XTIC,YMIN.P+TICY),1
2600 LINE (XTIC+XP.P,YMIN.P)-(XTIC+XP.P,YMIN.P+TICY),1
2610 LINE (XTIC,YMAX.P-TICY)-(XTIC,YMAX.P),1
2620 LINE (XTIC+XP.P,YMAX.P-TICY)-(XTIC+XP.P,YMAX.P),1
2630 XPOS.P=((XLOW.P+(XTIC-XMIN.P))/XT.P)*80!)-(LEN(STR$(XTIC))/2)
2640 LOCATE 23,XPOS.P:PRINT USING "##.#";XTIC;
2650 NEXT XPOS.P
2660 NEXT YTIC
2670 'tic marks and numbers on linear y axis
2680 IF YLIN=1 THEN 2780
2690 FOR YTIC=YMIN.P TO YMAX.P STEP YINC.P
2700 LINE (XMIN.P,YTIC)-(XMIN.P+TIXC,YTIC),1
2710 LINE (XMAX.P-TIXC,YTIC)-(XMAX.P,YTIC),1
2720 YPOS.P=((YHI.P+(YMAX.P-YTIC))/YT.P)*26!
2730 XPOS.P=6-(LEN(STR$(XTIC))/2)
2740 IF YPOS.P>25 OR YPOS.P<1 THEN BEEP:GOTO 2800
2750 IF YPOS.P>80 OR YPOS.P<1 THEN BEEP:GOTO 2800
2760 LOCATE YPOS.P,XPOS.P:PRINT USING "##.#";YTIC
2770 NEXT YPOS.P
2780 NEXT YTIC
2790 '
tic marks and numbers on log x axis

IF XLIN=0 THEN 3010
FOR CYC=-5 TO 5
XTIC=LTIC*(10^CYC)
LXTIC=LOG(XTIC)/LOG(10)
IF LXTIC<XMIN.P OR LXTIC=XMAX.P THEN 2950
LINE (LXTIC,YMIN.P)-(LXTIC,YMIN.P+TICY),1
LINE (LXTIC+XP.P,YMIN.P)-(LXTIC+XP.P,YMIN.P+TICY),1
LINE (LXTIC,YMAX.P-TICY)-(LXTIC,YMAX.P),1
LINE (LXTIC+XP.P,YMAX.P-TICY)-(LXTIC+XP.P,YMAX.P),1
NEXT LTIC

IF LXTIC>=XMIN.P AND LXTIC<=XMAX.P THEN LOCATE 23,(((XLOW.P+(LXTIC-XMIN.P))/XT.P)*80!)-1:PRINT XTIC;
NEXT CYC

tic marks and numbers on log y axis

IF YLIN=0 THEN 3140
FOR CYC=-5 TO 5
YTIC=LTIC*(10^CYC)
LYTIC=LOG(YTIC)/LOG(10)
IF LYTIC<YMIN.P OR LYTIC=YMAX.P THEN 3090
LINE (XMIN.P,LYTIC)-(XMIN.P+TICX,LYTIC),1
LINE (XMAX.P-TICX,LYTIC)-(XMAX.P,LYTIC),1
NEXT LTIC

YPOS.P=((YHI.P+(YMAX.P-LYTIC))/YT.P)*26!
XPOS.P=6-((LEN(STR$(YTIC))/2!))
IF LYTIC>=YMIN.P AND LYTIC<=YMAX.P AND YPOS.P>=1 THEN LOCATE YPOS.P,XPOS.P;PRINT YTIC;
NEXT CYC

now plot data on axes

SX=0:SY=0:SSX=0:SXY=0
FOR I=1 TO NPTS
X1(I)=X(I):IF XLIN=1 THEN X1(I)=LOG(X(I))/LOG(10)
Y1(I)=Y(I):IF YLIN=1 THEN Y1(I)=LOG(Y(I))/LOG(10)
IF I>1 THEN LINE(X1(I-1),Y1(I-1))-(X1(I),Y1(I)),1,STYLE
IF I>1 THEN LINE (X1(I-1)+XP.P,Y1(I-1))-(X1(I)+XP.P,Y1(I)),1,STYLE
IF SYM=1 THEN LINE (X1(I)-DX,Y1(I)-DY)-(X1(I)+DX,Y1(I)+DY),1,B
IF SYM=1 OR SYM=2 THEN LINE (X1(I)-DX+XP.P,Y1(I)-DY)-(X1(I)+DX+XP.P,Y1(I)+DY),1,B
IF SYM=2 THEN LINE(X1(I)-DX,Y1(I)-DY)-(X1(I)+DX,Y1(I)+DY),1,BF
IF SYM=3 OR SYM=4 THEN LINE (X1(I)-DX,Y1(I)-DY)-(X1(I)+DX,Y1(I)-DY),1:LINE (X1(I),Y1(I)+DY)-(X1(I)-DX,Y1(I)+DY),1
IF SYM=3 OR SYM=4 THEN LINE (X1(I)+DX,Y1(I)-DY)-(X1(I)-DX,Y1(I)+DY),1:LINE (X1(I)-DY,Y1(I)+DX)-(X1(I)-DY,Y1(I)-DX),1
IF SYM=3 OR SYM=4 THEN LINE (X1(I)-DX+XP.P,Y1(I)-DY)-(X1(I)+DX+XP.P,Y1(I)+DY),1,BF
IF SYM=3 OR SYM=4 THEN LINE (X1(I)+DX+XP.P,Y1(I)-DY)-(X1(I)-DX+XP.P,Y1(I)+DY),1:LINE (X1(I)-DY,Y1(I)+DX)-(X1(I)-DY,Y1(I)-DX),1
DX
IF SYM=3 OR SYM=4 THEN LINE (X1(I)+DX,Y1(I)-DY)-(X1(I)-DX,Y1(I)+DY),1:LINE (X1(I)+DX,Y1(I)-DY)-(X1(I)-DX,Y1(I)+DY),1
IF SYM=5 OR SYM=6 THEN CIRCLE (X1(I),Y1(I)),DX:CIRCLE (X1(I)+DX,Y1(I)),DX
IF SYM=5 OR SYM=6 THEN CIRCLE (X1(I)-DX,Y1(I)),DX:CIRCLE (X1(I)-DX-Y1(I)),DX
IF SYM=5 OR SYM=6 THEN CIRCLE (X1(I)+DX-Y1(I)),DX:CIRCLE (X1(I)+DX,Y1(I)),DX
139
3330 IF SYM=8 THEN PAINT (X1(I)+2*XP.P,Y1(I)),1
3340 SY=SY+Y1(I):SX=SX+X1(I):SSX=SSX+(X1(I)^2):SXY=SXY+(X1(I)*Y1(I))
3350 NEXT I
3360 IF LTYPE<>4 THEN RETURN
3370 ' Regression line plotted
3380 A=((NPTS*SXY)-(SX*SY))/((NPTS*SSX)-(SX*SX))
3390 B=(SY/NPTS)-(A*SX/NPTS)
3410 LINE (XMIN.P,YMIN.P)-(XMAX.P,YMAX.P),1
3420 LINE (XMIN.P+XP.P,YMIN.P)-(XMAX.P+XP.P,YMAX.P),1
3430 ' key trap of Alt-C to change color
3440 ' KLR.P=(KLR.P+1) MOD 128:IF KLR.P MOD 8=0 OR KLR.P MOD 16=0 THEN KLR.P=KLR.P +1
3450 OUT 985,KLR.P
3460 RETURN
3470 OPEN "com1:9600,s,7,1,rs,cs65535,ds,CD" AS #1
3480 PRINT #1, "IN;SP1;IP1750,1100,8500,7550;"
3490 PRINT #1, "SCXMIN,XMAX,YMIN,YMAX;"
3500 PRINT #1, "PUXMIN,YMIN, PDXMAX,YMIN,XMAX,YMAX, XMIN,YMIN PU"
3510 PRINT #1, "ST.2,.3;TL1.5,0"
3520 I=0
3530 FOR XTIC=XMIN TO XMAX STEP XINC
3540 PRINT #1,"PA";X(I),",0":XT;
3550 PRINT #1,"CP-2,-1;LB";X(I),+CHR$(3)
3560 I=I+1:NEXT XTIC

140
Appendix 2. Explanation of symbols for each mango sample.

<table>
<thead>
<tr>
<th>Tommy Atkins Samples</th>
<th>Haden Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>tali1mil</td>
<td>h11mil</td>
</tr>
<tr>
<td>tali1m12</td>
<td>h11m12</td>
</tr>
<tr>
<td>tali1m13</td>
<td>h11m13</td>
</tr>
<tr>
<td>tali1m14</td>
<td>h11m14</td>
</tr>
<tr>
<td>tali1m15</td>
<td>h11m15</td>
</tr>
<tr>
<td>tali1mm1</td>
<td>h11mm1</td>
</tr>
<tr>
<td>tali1mm2</td>
<td>h11mm2</td>
</tr>
<tr>
<td>tali1g11</td>
<td>h11g11</td>
</tr>
<tr>
<td>tali1g12</td>
<td>h11g12</td>
</tr>
<tr>
<td>tali1g13</td>
<td>h11g13</td>
</tr>
<tr>
<td>tali1g14</td>
<td>h11g14</td>
</tr>
<tr>
<td>tali1g15</td>
<td>h11g15</td>
</tr>
<tr>
<td>tali1gm1</td>
<td>h11gm1</td>
</tr>
<tr>
<td>tali1gm2</td>
<td>h11gm2</td>
</tr>
<tr>
<td>tali2m11</td>
<td>h12m11</td>
</tr>
<tr>
<td>tali2m12</td>
<td>h12m12</td>
</tr>
<tr>
<td>tali2m13</td>
<td>h12m13</td>
</tr>
<tr>
<td>tali2m14</td>
<td>h12m14</td>
</tr>
<tr>
<td>tali2m15</td>
<td>h12m15</td>
</tr>
<tr>
<td>tali2mm1</td>
<td>h12mm1</td>
</tr>
<tr>
<td>tali2mm2</td>
<td>h12mm2</td>
</tr>
<tr>
<td>tali2g11</td>
<td>h12g11</td>
</tr>
<tr>
<td>tali2g12</td>
<td>h12g12</td>
</tr>
<tr>
<td>tali2g13</td>
<td>h12g13</td>
</tr>
<tr>
<td>tali2g14</td>
<td>h12g14</td>
</tr>
<tr>
<td>tali2g15</td>
<td>h12g15</td>
</tr>
<tr>
<td>tali2gm1</td>
<td>h12gm1</td>
</tr>
<tr>
<td>tali2gm2</td>
<td>h12gm2</td>
</tr>
<tr>
<td>tali3m11</td>
<td>h13m11</td>
</tr>
<tr>
<td>tali3m12</td>
<td>h13m12</td>
</tr>
<tr>
<td>tali3m13</td>
<td>h13m13</td>
</tr>
<tr>
<td>tali3m14</td>
<td>h13m14</td>
</tr>
<tr>
<td>tali3m15</td>
<td>h13m15</td>
</tr>
<tr>
<td>tali3mm1</td>
<td>h13mm1</td>
</tr>
<tr>
<td>tali3mm2</td>
<td>h13mm2</td>
</tr>
<tr>
<td>tali3g11</td>
<td>h13g11</td>
</tr>
<tr>
<td>tali3g12</td>
<td>h13g12</td>
</tr>
<tr>
<td>tali3g13</td>
<td>h13g13</td>
</tr>
<tr>
<td>tali3g14</td>
<td>h13g14</td>
</tr>
<tr>
<td>tali3g15</td>
<td>h13g15</td>
</tr>
<tr>
<td>tali3gm1</td>
<td>h13gm1</td>
</tr>
<tr>
<td>tali3gm2</td>
<td>h13gm2</td>
</tr>
<tr>
<td>tali4m11</td>
<td>h14m11</td>
</tr>
<tr>
<td>tali4m12</td>
<td>h14m12</td>
</tr>
<tr>
<td>tali4m13</td>
<td>h14m13</td>
</tr>
<tr>
<td>tali4m14</td>
<td>h14m14</td>
</tr>
<tr>
<td>tali4m15</td>
<td>h14m15</td>
</tr>
<tr>
<td>tali4mm1</td>
<td>h14mm1</td>
</tr>
<tr>
<td>tali4mm2</td>
<td>h14mm2</td>
</tr>
<tr>
<td>tali4g11</td>
<td>h14g11</td>
</tr>
<tr>
<td>tali4g12</td>
<td>h14g12</td>
</tr>
<tr>
<td>tali4g13</td>
<td>h14g13</td>
</tr>
<tr>
<td>tali4g14</td>
<td>h14g14</td>
</tr>
<tr>
<td>tali4g15</td>
<td>h14g15</td>
</tr>
<tr>
<td>tali4gm1</td>
<td>h14gm1</td>
</tr>
<tr>
<td>tali4gm2</td>
<td>h14gm2</td>
</tr>
</tbody>
</table>

11-14 = lot numbers  
i= individual samples  
m= first m= ripened on the tree  
m= second m= mixed samples  
g= ripened in storage
Appendix 3. Sample calculation for the Box plot

Intensity: treatment #1

<table>
<thead>
<tr>
<th>Raw data</th>
<th>Raw data reorganized in ascending order</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.90</td>
<td>1.95 - whisker</td>
</tr>
<tr>
<td>6.97</td>
<td>2.31</td>
</tr>
<tr>
<td>9.98</td>
<td>3.13</td>
</tr>
<tr>
<td>2.31</td>
<td>3.61</td>
</tr>
<tr>
<td>6.14</td>
<td>3.90</td>
</tr>
<tr>
<td>4.69</td>
<td>4.07 - lower hinge</td>
</tr>
<tr>
<td>5.21</td>
<td>4.12</td>
</tr>
<tr>
<td>6.13</td>
<td>4.64</td>
</tr>
<tr>
<td>7.05</td>
<td>4.69</td>
</tr>
<tr>
<td>4.07</td>
<td>5.21 - Median</td>
</tr>
<tr>
<td>5.54</td>
<td>5.54</td>
</tr>
<tr>
<td>3.13</td>
<td>6.13</td>
</tr>
<tr>
<td>6.51</td>
<td>6.14</td>
</tr>
<tr>
<td>1.95</td>
<td>6.14</td>
</tr>
<tr>
<td>7.38</td>
<td>6.51 - upper hinge</td>
</tr>
<tr>
<td>6.14</td>
<td>6.97</td>
</tr>
<tr>
<td>8.07</td>
<td>7.05</td>
</tr>
<tr>
<td>4.64</td>
<td>7.38</td>
</tr>
<tr>
<td>4.12</td>
<td>8.07</td>
</tr>
<tr>
<td>3.61</td>
<td>9.98 - whisker</td>
</tr>
</tbody>
</table>

Equations used:

\[
\text{Median location} = \frac{\# \text{ of observations} + 1}{2} = \frac{20 + 1}{2} = 10.5
\]

\[
\text{Median} = \frac{\# \text{ at median location} \text{ (if odd \#, use the average of the two middle \#'s)}}{2}
\]

\[
= \frac{5.21 + 5.54}{2} = 5.38
\]

142
Equations cont.

Quartile location = \frac{\# \text{ of observations}}{2} + 1 = \frac{20}{2} + 1 = \frac{20}{2} + 1 = 5.5

Lower hinge = \# \text{ at quartile location (from top)} = 5.5 \approx 6 = 4.07

Upper hinge = \# \text{ at quartile location (from bottom)} = 6.51

Whiskers = \text{high and low range of observations} = 1.95 \& 9.98