EVALUATION OF A VAPOR-LIQUID-EQUILIBRIUM DATABASE USING ERROR ANALYSES BASED ON AN EQUATION OF STATE AND STATISTICAL METHODS

By

SZE-HOONG WONG

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

1997

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 1999

EVALUATION OF A VAPOR-LIQUID-EQUILIBRIUM DATABASE USING ERROR ANALYSES BASED ON AN EQUATION OF STATE AND STATISTICAL METHODS

Thesis Approved:

Thesis Adviser

Arland H. Johannes

K.A.M. Gasem

Wayne B. Powell

Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to take this golden opportunity to express my gratefulness to my mentor and thesis adviser, Dr. Jan Wagner, for his guidance and valuable advice. There is no doubt of my privilege and honor to have worked under his supervision. My special thanks also go to Dr. K. A. M. Gasem and Dr. A. Johannes, the second and third members of my thesis committee, for their recommendations, revisions, and critical assessment of this work.

For the financial support of this work, I would like to thank Gas Processors Association, Tulsa, Oklahoma. It was really a great financial relief with the financial help of this organization.

I continue express my gratitude to my only one God, Jesus Christ, for giving me intelligence and faith for finishing this remarkable work. It was an incredible joyful experience to have Jesus' accompaniment throughout this work.

Finally, I would like to express my indebtedness to my family in Malaysia, my lovely girl friend, and friends in the United States for their love, encouragement, understanding and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

English Letters

a	Parameter in the SRK EOS, Equation (3-1)
A, B, D	Adjustable parameters for four-suffix Margules equation
A, B	Parameters in Equation (3-8)
A'', B''	Parameters in Equation (3.13)
b	Parameter in the SRK EOS, Equation (3-1)
exp	Exponential
EOS	Equation of state
F	Fisher-ratio
G^E	Excess gibbs free energy
GPA	Gas Processors Association
ΔH	Molar enthalpy of mixing
H_o	Hypothesis
k	Binary interaction parameter
K	K-value
$\delta \ln(f_1 / f_2)$	Residual term derived from Equation (2.19)
$\bar{P}$	Sample mean of the variable P
$\hat{P}$	SRK model predictions of variable P

P	Absolute pressure
R	Universal gas constant
RMSE	Root mean square error
$S^2(P)$	Estimated variance of the variable P
$S^2(\hat{P})$	Residual variation from the line of model predictions
$S^2(c)$	Variance of the model predictions
T	Absolute temperature
Δ V	Molar volume of mixing
v	Volume
VLE	Vapor-liquid-equilibrium
x	Component mole fraction in the liquid phase
y	Component mole fraction in the vapor phase
Z	Compressibility factor

Subscripts

c	Physical properties at critical state
i	Number of component
l	Liquid phase
v	Vapor phase

Superscripts

expt	Experimental data
o	Standard condition
sat	Saturated properties

Greek Symbols

γ	Activity coefficient
ω	Accentric factor
ϕ	Fugacity coefficient
δ	Residual
α	Temperature dependent parameter, used in the SRK EOS
ε	Simplified term in Equation (2.19)
Σ	Summation

CHAPTER I

INTRODUCTION

On August 1, 1993, two projects were initiated under a single contract between the Gas Processors Association (GPA) and Oklahoma State University (OSU): GPA Project 921, Ethalphy Database and Maintenance, and GPA Project 925, Maintenance and Evaluation of Data for the GPA Data Bank. This document is a continued work of GPA Project 925 and mainly concentrates on the evaluation of vapor-liquid-equilibrium (VLE) data.

According to the database summary statistics, there are approximately 16,000 VLE data records, about 37 percent of the data records in GPA data bank. These data records mainly include the measurements of pressure, temperature, component mole fraction in the liquid phase, and component mole fraction in the vapor phase. About 13 percent of the VLE data records include the component mole fractions in feed.

These VLE properties are particularly important for the design of process equipment (such as distillation columns, absorbers, flash separators and heat exchangers) and for the development and evaluation of equations of state. Therefore, it is necessary to detect both data entry errors and the systematic errors, deviations attributed to experimental measurements, in the VLE data.

Two approaches can be used for this study. They are:

(i) Thermodynamic consistency tests based on Gibbs-Duhem equation. Procedures reported in the literature are classified broadly into point tests (4, 6, 11, 12, 24, 28, 38, 33) and area tests (2, 8, 9, 20, 22, 23). Other well-known methods are the data reduction method (32) and the direct test method (31).

(ii) Error analyses based on an equation-of-state (EOS) model and statistical methods. These analyses include three methodologies. The first and second methodologies are based on a statistical method – the Fisher-ratio test (F-test). The Soave-Redlich-Kwong (SRK) EOS is used in the third methodology. This EOS is recommended, since most of the VLE systems in GPA Database contain light hydrocarbon components.

Only a limited number of VLE systems in the GPA Database can be tested by thermodynamic consistency tests based on Gibbs-Duhem equation, since the experimental volumetric and enthalpy data are not available at the required temperature and pressure conditions. Thus, error analyses based on SRK EOS and F-tests were employed.

Research Objectives

The objectives of this study are:

(i) To identify the data sets that are amenable to the error analyses approach and to carry out the appropriate tests.

(ii) To evaluate the use of statistical methods in screening the VLE data.

(iii) To use the SRK EOS and statistical methods to assess the quality of the VLE data in the GPA databank.

This document is organized into six chapters. Chapter II presents the literature review of the thermodynamic consistency tests and the error analyses of this study. As an overview of the first approach, common methods and definition of thermodynamic consistency tests are discussed. For the second approach, definition of an equation of state is given. A general description of the cubic EOS is presented, and it includes the reasons for choosing the SRK EOS as a model to predict the VLE data. Also, the applications of the F-tests are reviewed to determine whether

(i) The same component data sets reported by different investigators came from the same population.

(ii) The model predictions are reasonable.

A three-step procedure that is used to identify test cases from the GPA Database is discussed in Chapter III.

Chapter IV presents the basic equations of SRK EOS and a general discussion of bubble-point calculations. A Fortran program used to model the desired VLE data is also described.

Results and discussions of the VLE data evaluation are given in Chapter V. It includes three methodologies, which are demonstrated with sample cases, to evaluate the use of statistical methods in screening the VLE data in the GPA Database. Also, data discrepancies and data meriting further examination were identified. Finally, conclusions and recommendations of this study can be found in Chapter VI.

CHAPTER II

LITERATURE REVIEW

This chapter presents a discussion of the background concepts of thermodynamic consistency tests and the error analyses based on Soave-Redlich-Kwong (SRK) equation of state (EOS) and F-tests. Definitions of thermodynamic consistency tests and EOS are included. A summary of the common methods for these two approaches is given, including their advantages and limitations.

Approach 1: Thermodynamic Consistency Test Based on Gibbs-Duhem Equation

Definition of Thermodynamic Consistency Test

A thermodynamic consistency test is a check for internal consistency of the experimental VLE data by means of basic thermodynamic equations. The Gibbs-Duhem equation (either in its differential or integral form) is the most widely used (1). Common methods that have applied Gibbs-Duhem equation are point tests, area tests, data reduction method and direct test method. These methods are discussed in the following sections.

Application of Point Test (Slope Test)

Point tests are usually used to detect serious errors in the experimental VLE data. For a binary system, the general form of Gibbs-Duhem equation can be written as follows (19):

$$x_1 \frac{d \ln \gamma_1}{dx_1} + x_2 \frac{d \ln \gamma_2}{dx_2} = -\frac{\Delta V}{RT} dP + \frac{\Delta H}{RT^2} dT \tag{2.1}$$

where ΔV and ΔH represent the molar volume of mixing and molar enthalpy of mixing, respectively. Also, by definition, experimental values of the activity coefficient of species 1 and 2 can be calculated using Equation (2.2):

$$\gamma_i = \left(\frac{y_i P}{x_i P_i^{sat}} \right) \phi_i, \quad \text{(i=1, 2)} \tag{2.2}$$

where the Poynting correction (19) is expressed as

$$\phi_i = \left(\frac{\hat{\phi}_i}{\phi_i^{sat}} \right) \exp\left\{ \frac{V_i^l \left(P - P_i^{sat} \right)}{RT} \right\} \tag{2.2a}$$

The saturated vapor pressure, P_i^{sat}, is a strong function of temperature. At low and moderate pressures, Poynting correction is considered negligible.

At constant pressure and temperature, Equation (2.1) can be simplified as follows (19):

$$x_1 \frac{d \ln \gamma_1}{dx_1} + x_2 \frac{d \ln \gamma_2}{dx_2} = 0 \tag{2.3}$$

This equation is used to test the experimental VLE data directly. Plots of $\ln \gamma_1$ as a function of x_1 and $\ln \gamma_2$ as a function of x_2 are prepared and slopes are measured (19).

Values of the slopes are substituted into Equation (2.3) at various compositions to see if the Gibbs-Duhem equation is satisfied.

The limitation of this method is that it is difficult to measure the slopes with sufficient accuracy. This method requires a large number of experimental data to obtain accurate values of the slopes. Also, most of the VLE data in the GPA Database are not measured at constant pressure and constant temperature, so Equation (2.3), which is over-simplified, is not recommended for this application.

Application of Area Test

For quantitative purposes, area tests are much easier to use compared to the slope tests. Most of the integral test equations were discussed exhaustively by Redlich and Kister (22) and Herington (9). Conceptually, the integral form of the Gibbs-Duhem equation is (37)

$$\sum x_i d\ln\gamma_i = -\int_{P_2}^{P_1}\frac{\Delta V}{RT}dP + \int_{T_2}^{T_1}\frac{\Delta H}{RT^2}dT \tag{2.4}$$

This equation is generally applied to the VLE data that are not at constant temperature and pressure conditions.

For isothermal and isobaric systems, Equation (2.4) becomes (38)

$$\int_0^1 \ln\frac{\gamma_1}{\gamma_2}dx_1 = 0 \tag{2.5}$$

Equation (2.5) is called the area test for the phase equilibrium data (19). A plot of $\ln(\gamma_1/\gamma_2)$ as a function of x_1 is prepared and thermodynamic consistency is met when the net area under the curve is zero.

To calculate the ratio of γ_1 and γ_2, Equation (2.2) can be applied. This can be shown as follows in which the pressure cancels out (19):

$$\frac{\gamma_1}{\gamma_2} = \frac{(\varphi_1 y_1 P)/(x_1 P_1^{sat})}{(\varphi_2 y_2 P)/(x_2 P_2^{sat})} = \frac{(\varphi_1 y_1)/(x_1 P_1^{sat})}{(\varphi_2 y_2)/(x_2 P_2^{sat})} \tag{2.6}$$

The cancellation of the pressure, which is the most accurate measurement, is considered a major disadvantage of area test.

According to Van Ness (32), the area test generally uses the x-y values and the ratio of the two pure-component vapor pressures, P_1^{sat} / P_2^{sat} to construct the plot. In an isothermal system, "this method does little more than determine whether or not the vapor pressure ratio P_1^{sat} / P_2^{sat} is appropriate to the set of measured x-y value" (19). Therefore, it tells us nothing about the internal consistency of the VLE data, when we make a plot of $\ln(\gamma_1 / \gamma_2)$ as a function of x_1.

The last problem associated with this method arises from the VLE results of isobaric or isothermal systems. For isobaric systems, the following equation is derived from Equation (2.4):

$$\int_0^1 \ln\frac{\gamma_1}{\gamma_2}dx_1 = \int_{T_2}^{T_1} \frac{\Delta H}{RT^2}dT \tag{2.7}$$

The right hand side of this equation is a quantity that cannot be generally ignored (37). However, the required enthalpy values are often missing, so the right hand side term is considered negligible and omitted. This may introduce significant errors in the consistency test. For isothermal cases, the main equation is

$$\int_0^1 \ln\frac{\gamma_1}{\gamma_2}dx_1 = \int_{P_2}^{P_1} \frac{\Delta V}{RT}dP \tag{2.8}$$

This equation can produce an excellent approximation for the consistency test even though the information of the molar volume of mixing is not available. The main reason is that the right hand side of Equation (2.8) is often very small at low or moderate pressure. However, for the high-pressure isothermal cases, the computations of the right hand side term are necessary.

Application of Data Reduction Method

Since experimental values of *P*, *T*, *x*, *y* are commonly found in VLE systems, Van Ness and coworkers (32) proposed a data reduction method. They claimed that this method is a more meaningful check for the thermodynamic consistency of VLE data. Their ideas were focused on overcoming the problems and limitations of area tests.

The pertinent equations of this method, which are restricted to binary systems at low pressure and constant temperature, are (32)

$$\gamma_i = \frac{y_i P}{x_i P_i^{sat}} \tag{2.9}$$

$$\frac{G^E / RT}{x_1 x_2} = \frac{\ln \gamma_1}{x_2} + \frac{\ln \gamma_2}{x_1} \tag{2.10}$$

where

$$\ln \gamma_1 = \frac{G^E}{RT} + x_2 \frac{d\left(G^E / RT\right)}{dx_1} \tag{2.11a}$$

and

$$\ln \gamma_2 = \frac{G^E}{RT} + x_1 \frac{d\left(G^E / RT\right)}{dx_2} \tag{2.11b}$$

By manipulating the Equations (2.9) through (2.11), they become (32)

$$\ln\frac{\gamma_1}{\gamma_2} = \ln\frac{y_1 x_2 P_2^{sat}}{y_2 x_1 P_1^{sat}} = \frac{d\left(G^E / RT\right)}{dx_1} \tag{2.12}$$

In Equation (2.9), the vapor phase fugacity coefficient is assumed negligible at low and moderate pressures. However, for high-pressure systems, Equation (2.2) is recommended. The liquid phase properties are assumed to be independent of pressure as shown in Equation (2.9).

For the majority of systems, the empirical functional relation between G^E / RT and x can be calculated by the following empirical, 3-parameter Margules equation (32):

$$\frac{G^E}{RT} = \left(Ax_2 + Bx_1 - Dx_1x_2\right)\left(x_1x_2\right) \tag{2.13}$$

where the parameters A, B, and D are functions of temperature.

The following equations, which are often called four-suffix Margules equations, can be derived by substituting Equation (2.13) into Equations (2.11) and (2.12) (32):

$$\ln\gamma_1 = x_2^2\left[A + 2(B - A - D)x_1 + 3Dx_1^2\right] \tag{2.14a}$$

$$\ln\gamma_2 = x_1^2\left[B + 2(A - B - D)x_2 + 3Dx_2^2\right] \tag{2.14b}$$

$$\ln\frac{\gamma_1}{\gamma_2} = Ax_2^2 - Bx_1^2 + 2x_1x_2\left[B - A + D(x_1 - x_2)\right] \tag{2.15}$$

To apply the thermodynamic consistency test, there are two different approaches. The most direct approach is to calculate γ_1 and γ_2 using Equation (2.2) for each data point. Then, Equation (2.10) is used to generate a set of values for G^E / RT. The values of G^E / RT are fit by a least-squares technique to Equation (2.13) (32).

The second approach is divided into three distinctly different procedures. Each makes use of data for just two of the variables--x-y, P-x, or P-y in vapor liquid equilibrium (32). According to Barker's method (1), the liquid composition is assumed to not be affected by experimental errors. Therefore, calculated values of y can then be compared with measured values as a check on the thermodynamic consistency of the data (32).

The following methods require only P-x or P-y data (32):

$$\frac{dy_1}{dP} = \frac{y_1 y_2}{P(y_1 - x_1)} \tag{2.16a}$$

This is a form of coexistence equation, and it can be solved for x_1:

$$x_1 = y_1\left(1 - \frac{y_2}{P}\frac{dP}{dy_1}\right) \tag{2.16b}$$

For binary data, an additional equation can be derived directly from Equation (2.9):

$$P = x_1 P_1^{sat}\gamma_1 + x_2 P_2^{sat}\gamma_2 \tag{2.17a}$$

Substituting Equation (2.11) into Equation (2.17a),

$$\begin{aligned} P = x_1 P_1^{sat} \exp\left[\frac{G^E}{RT} + x_2 \frac{d\left(G^E / RT\right)}{dx_1}\right] \\ + x_2 P_2^{sat} \exp\left[\frac{G^E}{RT} + x_1 \frac{d\left(G^E / RT\right)}{dx_1}\right] \end{aligned} \tag{2.17b}$$

Since the liquid composition is assumed not to be affected by experimental errors, P-x data are used. Equation (2.16a) or Equation (2.17b) is therefore implemented. Equation (2.16a) can be integrated to yield y-values. Activity coefficients and G^E / RT or $\ln(\gamma_1 / \gamma_2)$ are calculated with the measured P and x data. With Equation (2.17b), we can use the numerical method proposed by Mixon (15) to search for the relation between

G^E/RT and x (32). This will allow Equation (2.17b) to reproduce the measured P as a function of x data as closely as possible.

Application of Direct Test

After the publication of the data reduction method, Van Ness (31) provided another method far more powerful than the area test-- the direct test. This method applies the concepts of the data reduction method and area test, which are illustrated in the following paragraphs. By applying this method, the deviations of the VLE data from the Gibbs-Duhem equation can be evaluated directly.

The general equation of area test is the same as Equation (2.4). By the definition of Van Ness (31), the area test equation that describes binary VLE systems at constant pressure or constant temperature condition can be expressed as:

$$\int_0^1 \left\{ \ln\left(\frac{\gamma_1^{\exp t}}{\gamma_2^{\exp t}}\right) + \varepsilon \right\} dx_1 = 0 \tag{2.18}$$

For simplicity, the definitions of ε are invoked from Equation (2.4):

$$\varepsilon = \varepsilon_T = -\left(\frac{\Delta H}{RT^2}\right)\frac{dT}{dx_1} \tag{2.18a}$$

for isobaric cases. For isothermal cases,

$$\varepsilon = \varepsilon_P = -\left(\frac{\Delta V}{RT}\right)\frac{dP}{dx_1} \tag{2.18b}$$

In either case, only one ε term is needed. The ratio of $\gamma_1^{\exp t}/\gamma_2^{\exp t}$ is calculated using Equation (2.6) in which the experimental measurements of P, T, x, and y are used.

Now, the general form of the direct test is (31)

$$\int_0^1 \left[\ln\left(\frac{\gamma_1}{\gamma_2}\right) - \ln\left(\frac{\gamma_1^{\exp t}}{\gamma_2^{\exp t}}\right)\right] dx_1 = -\int_0^1 \left\{\ln\left(\frac{\gamma_1^{\exp t}}{\gamma_2^{\exp t}}\right) - \varepsilon\right\} dx_1 \tag{2.19}$$

For simplicity, this equation can be written as (31):

$$\int_0^1 \delta \ln\left(\frac{f_1}{f_2}\right) dx_1 = -\int_0^1 \left\{\ln\left(\frac{\gamma_1^{\exp t}}{\gamma_2^{\exp t}}\right) - \varepsilon\right\} dx_1 \tag{2.19a}$$

where $\delta \ln(f_1 / f_2)$ is the residual term on the left hand side of Equation (2.19). The ratio of γ_1 / γ_2 can also be calculated using Equation (2.6). However, the pressure and vapor composition values in Equation (2.6) are provided by a pressure dependent bubble-point calculation in which temperature and the liquid composition values are given (31). Thus, the residual term is defined as the difference between a derived value and corresponding experimental value.

In Equation (2.19), the negative sign before the integration on the right hand side is needed due to the derivation of $\ln\left(\gamma_1^{\exp t} / \gamma_2^{\exp t}\right)$. Thus, the right hand side of this equation is exactly the quantity given by Equation (2.18). The residual on the left plays an important role as a direct measure of deviations from the Gibbs-Duhem equation in terms of area test. In other words, a plot of these residuals against x_1 for the direct tests also displays the area tests (30). However, for this specific purpose a more suitable objective function is $\sum \{\delta \ln(f_1 / f_2)\}^2$. This objective function causes the residuals to scatter about a horizontal line.

Although this method is unique for the consistency test and has more features than the area test, it poses the same disadvantages as in the area test. This method does not utilize the more accurate pressure measurements to calculate the ratio of γ_1 / γ_2 or $\gamma_1^{\exp'} / \gamma_2^{\exp'}$. Also, the data for calculating the ε term are often missing.

Summary of Thermodynamic Consistency Methods

The following table gives a brief summary of each method discussed above. It includes their advantages, limitations, and reasons for limited value in this work:

Table 2.1. Summary of Advantages, Limitations, and Applications for this work.

Methods	Advantages	Limitations	Comments Relevant to the VLE Database
Point Tests	Equations are easy to apply. They can detect serious errors of VLE data.	It is difficult to measure slopes with sufficient accuracy. They require an extensive compilation of VLE data.	Most VLE data sets do not have extensive compilation of data points.
Area Tests	For quantitative purposes, they are much easier to use compared to slope tests.	They do not utilize the total pressure measurement. They tell us nothing about the internal consistency of the VLE data when we make a plot of $\ln(\gamma_1/\gamma_2)$ versus x_1.	The required enthalpy and molar volume of mixing values are missing in the VLE database.
Data Reduction Method	It utilizes the total pressure measurement.	It is restricted to binary systems at constant temperature.	Limited number of data sets were measured in these conditions.
Direct Test	It provides a direct measure of the deviation from the Gibbs-Duhem equation.	It is restricted to binary systems. It does not utilize the total pressure measurement.	The required enthalpy and molar volume of mixing values are missing in the VLE database.

Approach 2: Error Analyses Based on an Equation-of-State Model and Statistical Methods

Definition of Equation of State (EOS)

By definition, an EOS is a mathematical relation among temperature, T; pressure, P; volume, v; and composition, x. It can be expressed as f (T, P, v, x) = 0, or in a pressure explicit form as P = f (T, v, x). For correlating thermodynamic properties, it provides the most efficient method. Also, the development of EOS have allowed us to seek models which are capable of describing the phase behavior of a variety of chemical species exhibiting varying degrees of shape-size effects, polarity, and association (7).

Groups of Equation of State (EOS)

There are four groups of EOS commonly found in the literature. They are (7):

(i) The van der Waals family of cubic equations, which are less accurate but widely used due to their simplicity and qualitative success in the industry.

(ii) The family of extended virial equations that are tailored to fulfill the need of a given industry.

(iii) Accurate substance-specific equations used to describe the thermodynamic properties of particular chemical species.

(iv) Equations evolved from advances in molecular thermodynamics, which are theoretically more rigorous.

Since SRK EOS is one of the van der Waals family of cubic equations, and the error analyses are based on this model; the discussions in the following section are focused on the cubic EOS.

Cubic Equation of State (EOS)

The van der Waals (vdW) equation was proposed in 1873 (7):

$$P = \frac{RT}{v-b} - \frac{a}{v^2} \tag{2.20}$$

where b is the excluded volume and a is the cohesive parameter. This simplified molecular model envisioned the system pressure as contributions by the molecular attraction effects and repulsive effects. The calculations of these effects are shown in the right hand side of Equation (2.20), respectively. However, this equation is not quantitatively accurate.

Numerous modifications in later model development have made the predictions more accurate. Three currently popular equations are particularly successful in improving the accuracy of van der Waal equation. They are the Redlich-Kwong (RK), the Soave-Redlich-Kwong (SRK), and the Peng-Robinson (PR) EOS (7). Many other cubic EOS have been proposed. Some are very complex and often better for specific applications (7).

The cubic EOS is one of the closed-form equations. The success of closed-form equations can be viewed from several advantages as follows (7):

(i) The same-form applies at all conditions and for different phases.

(ii) Direct solutions for volume can be obtained from some of the EOS, such as the cubic.

(iii) Critical and physical properties can be used to estimate EOS constants.

(iv) Strategies can be employed to estimate thermodynamic properties of non-model components, such as petroleum fractions.

(v) They provide a proven format for interpolating experimental data.

(vi) They are amenable to theoretical development.

However, deficiencies still exist among the cubic equations. Their limitations are as follows (7):

(i) They generally work well for normal fluids, but fail for highly polar, associating fluids and mixtures containing large molecules.

(ii) They work poorly near the critical point.

(iii) Only careful EOS tuning can obtain accurate predictions.

(iv) They are not suitable for extrapolations due to their empirical or semi-empirical nature.

(v) Generally, they predict volumetric properties poorly.

In this study, reasons for choosing SRK EOS to predict the VLE data can be viewed from the following advantages. Specifically, the SRK EOS is (18, 19):

(i) Applicable to non-polar or mildly polar mixtures,

(ii) Suitable in the high temperature and high-pressure regions,

(iii) Reasonably accurate if $T < 0.9\ T_c$, and therefore does not exhibit anomalous behavior,

(iv) Generalized and applicable to multi-component systems with established mixing rules, and

(v) Reasonably accurate with an acceptable speed of computation.

The basic equations of SRK EOS that were used in the third methodology are presented in Chapter IV of this document.

Statistical Methods

In the GPA Database, many VLE systems, especially binary VLE systems, are composed of data sets reported by different investigators. The measurements in these data sets are not expected to be identical, even though they came from the same VLE systems or the same population. The means or variances of the data sets in a VLE system will vary over a range which can be approximated for any desired probability levels. Statistical methods adopted in this study, F-tests, were applied as a quick screening tool to determine whether the ratio of two estimated variances is larger than might be expected by chance, if the tested data sets had been drawn from the same population.

This testing method is named the "population-check methodology". The null hypothesis of these variance tests, H_o, states that the variances of the data sets are equal to each other. The outcomes of this methodology indicate whether there are any major differences in the data measurements of the same-components data sets reported by different investigators.

The results of F-tests in this methodology are further analyzed by the second and third methodologies of the error analyses. If the evaluations of F-tests are shown to be reliable, VLE data sets in the GPA Database can be screened quickly to determine whether there are any major differences in the data sets without physically checking the individual data points.

In the second methodology, F-tests were used to determine whether model predictions are reasonable for a data set. This testing method is named the "model-predictions-check methodology".

Even though the measurement variables of the VLE data sets are properly related in the SRK EOS, there will almost certainly be some variations in the model predictions that cannot be modeled, or explained. These unexplained variances are assumed to be caused by the unexplainable random phenomena, so they can be referred to as random error (34).

To determine whether the explained variance is significant when compared to the unexplained variance, F-tests were used. These two types of estimated variances are further defined in the last section of this chapter. The null hypothesis of these variance tests states that the differences between these two estimated variances are significant when compared to each other.

The outcomes of F-tests in this methodology can be used to determine whether

(i) The SRK EOS predicts equally well in the data sets expected to come from the same population.

(ii) The SRK EOS can be used to model the VLE data sets in the GPA Database.

The common assumptions of F-tests in these two methodologies are (34):

(i) Both sampled populations are normally distributed.

(ii) The samples are random and independent.

Due to these assumptions, further information is needed to judge cases where the calculated F-ratios are close to the tabulated F-values.

The following sections present the equations and expected results of these two methodologies.

Population-Check Methodology

The main equation for the F-tests is (34):

$$F = \frac{\text{Larger Sample Variance}}{\text{Smaller Sample Variance}} = \frac{S^2\left(P'\right)}{S^2\left(P\right)} \tag{2.21}$$

where the estimated variance of the variable P is defined as the mean squared deviation from the sample mean, $\bar{P}$:

$$S^2(P) = \frac{\sum\left(P - \bar{P}\right)^2}{n-1} \tag{2.22}$$

To test the hypothesis that the sample variances are equal, one-sided tests are used. In these tests, the largest possible ratios are calculated using Equation (2.21) and compared to the tabulated F-values (34). These tabulated F-values are for the usual variance tests corresponding to the 0.05 and 0.01 probability, or significant, levels.

If the ratio of the estimated variances is smaller than the tabulated F value at the 0.05 or 0.01 probability level, the hypothesis that the variances are equal, is accepted. Therefore, the tested data sets are expected to come from the same population.

On the other hand, two-sided tests can be used to determine whether there is a significant difference between two estimated variances. In these tests, the indicated probability levels need to be doubled. The results of these tests indicate the percent chance of being error if the hypothesis that the differences between two estimated variances are not significant, is accepted.

In this study, one-sided tests were used for the population-check methodology. The results of this methodology are presented in Chapter V

Model-Predictions-Check Methodology

In the variance tests of this methodology, there are three main sources of variation that need to be considered. In statistical terms, they are called explained variance, unexplained variance, and total variance. Explained variance is the variation that can be explained by the model. Unexplained, or residual, variance is the difference between experimental data and model prediction. Total variance is the sum of explained and unexplained variances.

Equation (2.22), which has n-1 degrees of freedom, is used to calculate the total variance. However, to calculate the unexplained variance from the model predictions and experimental VLE data, the following equation is used (34):

$$S^2\left(\hat{P}\right) = \frac{\sum\left(P - \hat{P}\right)^2}{n-2} \tag{2.23}$$

where P is the experimental variable and $\hat{P}$ is the model prediction. Based on the equations of SRK EOS, it is difficult to determine the degrees of freedom. To be conservative, n-2 is considered as the closest and safest number to use according to Moser (17).

To calculate the explained variance, $S^2(c)$ is the difference between the total and unexplained variances, or

$$S^2(c) = \frac{\sum\left(P - \bar{P}\right)^2 - \sum\left(P - \hat{P}\right)^2}{(n-1)-(n-2)} = \sum\left(P - \bar{P}\right)^2 - \sum\left(P - \hat{P}\right)^2 \tag{2.24}$$

Then, the F-ratio can be determined by letting (34)

$$F = \frac{S^2(c)}{S^2\left(\hat{P}\right)} \tag{2.25}$$

To test the hypothesis that the differences between the explained and unexplained variances are significant, the ratio of these two estimated variances is compared to the tabulated F-values (34). If this ratio is larger than the tabulated F value at the 0.05 or 0.01 probability level, the hypothesis is accepted. Therefore, the model predictions are expected to be reasonable.

CHAPTER III

PRESELECTION OF TEST CASES

This chapter includes a discussion of a three-step procedure that identifies the VLE test cases from the GPA Database. Matrix systems that show the components in binary, ternary, and multi-component systems for VLE and dew-point and bubble-point data are given. A summary of all the test cases and a description of the extracted data files for the test cases are also included.

Step1: System Identification and Screening

The first step of this pre-selection process is to have a clear picture of the components available in the binary, ternary, and multi-component systems. These systems are restricted to VLE and dew-point and bubble-point data in the GPA Database. The procedure of identifying these systems can be found in Appendix A.

Six matrix systems were created and mapped into two different ways. As shown in Figures 1 (in the pocket) and 2, maps of the binary systems were drawn into triangular tables. The rows and columns are labeled with the components in the systems. The available binary systems are indicated by Xs in the cells.

COMPONENT	Methane	Propane	n-Butane	n-Pentane	n-Hexane	n-Heptane	Toluene	Carbon Dioxide	Hydrogen Sulfide	Hydrogen	Nitrogen
Methane											
Propane											
n-Butane	X										
n-Pentane	X										
n-Hexane	X										
n-Heptane	X										
Toluene	X										
Carbon Dioxide	X										
Hydrogen Sulfide											
Hydrogen		X									
Nitrogen								X	X		

Note: The Xs identify the binary dew-point and bubble-point systems

Figure 2. Binary Dew-Point and Bubble-Point Systems in the GPA Database

In Figures 3, 4 (in the pocket), 5 and 6, the rows and columns of each system are labeled differently. The rows are labeled as data-set-number (DSN); whereas the columns are labeled with the names of the components of the appropriate systems. Therefore, the Xs in a row identify the components in a particular data set. This mapping method can also be used to find the data sets that have the identical components.

Indirectly, these matrix systems provided the means to clearly identify the pure components physical properties which are required for the EOS model predictions.

Data Set Number (DSN)	COMPONENT						
	Methane	Ethane	Propane	Carbon Dioxide	Hydrogen sulfide	Nitrogen	Water
12	X	X		X			
13	X			X	X		
14	X			X		X	
15	X				X	X	
16	X			X	X		
5000	X		X				X

Note: The Xs in a row identify components in a data set of ternary dew-point and bubble-point system.

Figure 3. Ternary Dew-Point and Bubble-Point Systems in the GPA Database

Data Set Number (DSN)	COMPONENT																
	Methane	Ethane	Propane	Isobutane	n-Butane	n-Pentane	n-Hexane	n-Heptane	n-Octane	n-Decane	Ethene	Propene	Toluene	Carbon Dioxide	Hydrogen Sulfide	Hydrogen	Nitrogen
665	X	X														X	X
667	X	X	X		X	X											
668	X	X	X	X													X
669	X	X	X		X	X											X
670	X	X	X											X	X		X
671	X	X	X		X	X											X
672	X	X	X			X		X		X							
673	X	X	X		X	X											X
674	X	X	X			X		X		X			X				
675	X	X	X			X		X		X					X		
676	X	X	X			X		X		X			X		X		
677	X	X	X			X		X		X				X			X
678	X	X	X		X	X	X		X					X			X
679	X	X	X			X		X		X			X	X			X
680	X	X	X			X		X		X				X	X		X
764	X	X	X								X	X				X	
784	X	X	X			X	X										

Note: The Xs in a row identify components in a data set of multi-components VLE system.

Figure 5. Multi-Component Vapor-Liquid-Equilibrium Systems in the GPA Database

Data Set Number (DSN)	COMPONENT											
	Methane	Ethane	Propane	n-Butane	n-Pentane	n-Hexane	n-Heptane	n-Octane	Carbon Dioxide	Hydrogen Sulfide	Nitrogen	Water
17	X								X	X		X
18	X	X	X								X	
19	X	X	X	X	X						X	
20	X	X	X	X	X	X		X	X		X	
21	X	X	X	X	X	X	X	X	X		X	
758	X	X	X	X	X						X	
759	X	X	X	X	X						X	
760	X	X	X	X	X						X	

Note: The Xs in a row identify components in a data set of dew-point and bubble-point system.

Figure 6. Multi-Component Dew-Point and Bubble-Point Systems in the GPA Database

Step 2: Data Set Identification

To evaluate the use of the F-tests for screening the VLE data, same-component data sets reported by different investigators in the VLE systems were identified in the GPA Database. The data set numbers (DSN) and other important information of these data sets were found in the process of creating the matrix systems, which is documented in Appendix A.

Eighteen binary VLE systems that involved 42 data sets and one ternary VLE system that involved 2 data sets, were identified from the GPA Database. Summaries of these test cases are given in Tables 3.1 and 3.2. None of the dew-point and bubble-point systems are composed of data sets reported by different investigators, so this study focused only on evaluating the VLE systems.

Table 3.1. Test Cases in Binary VLE Systems

Test Case #	Components	Data Set #	#Pt	Authors	T(min) (R)	T(max) (R)	P(min) (psia)	P(max) (psia)
1	Methane	447	57	Kohn, J.P. & Shim, J.	536.67	873.27	0.2	2845.2
	n-Hexane	461	104	Chen, R.J.J., Chappelear, P.S. & Kobayashi, R.	342.90	491.69	19.9	2650.0
2	Methane	441	135	Kobayashi, R. & Wichterle, I.	234.67	359.87	0.2	748.0
	Ethane	505	27	Kobayashi, R., Wichterle, I. & Chappelear, P.S.	344.16	354.98	646.9	728.3
		529	25	Miller, R.C., Hiza, M.J. & Kidnay, A.J.	288.00	324.00	3.1	476.9
3	Methane	385	15	Kirk, B.S. & Ziegler, W.T.	185.37	209.67	293.9	1322.6
	Hydrogen	570	25	Masuoka, H., Yorizane, M., Toyama, A. & Yoshimura, S.	185.67	293.67	147.0	2204.4
4	Methane	428	29	Lacey, W.N. & Sage, B.H.	529.67	709.67	40.0	1900.0
	n-Butane	545	80	Wang, R.H., Azarnoosh, A., McKetta, J.J. & Roberts, L.R.	379.67	739.67	26.0	1915.0
		501	21	Sage, B.H., Jacobs, J. & Wiese, H.C.	499.67	679.67	200.0	1700.0
		587	4	Sauer, R.N.	559.67	559.67	111.0	1803.0
		539	7	Thodos, G., Rigas, T.J. & Mason, D.F.	559.67	559.67	960.0	1861.0
5	Ethane	450	15	Hoshino, D., Nagahama, K., Hirata, M. & Konishi, H.	455.31	455.31	207.2	335.1
	Carbon Dioxide	566	17	Hakuta, T., Nagahama, K. & Suda, S.	491.67	491.67	348.3	577.6
		519	15	Kidnay, A.J., Phelps, R.E., Davalos, J. & Anderson, W.R.	450.00	450.00	188.8	309.6
6	Ethane	423	27	Robinson, D.B., Krishnan, T.R. & Kalra, H.	359.87	509.67	14.3	433.2
	Hydrogen Sulfide	796	45	Kalra, H. & Robinson, D.B.	359.87	509.67	9.5	442.7
7	Propane	341	15	Akers, W.W., Lipscomb, T.G. & Kelley, R.E.	419.67	491.67	15.0	507.0
	Carbon Dioxide	451	24	Nagahama, K., Hirata, M., Hoshino, D. & Konishi, H.	455.31	491.67	35.3	505.5
		512	21	Lu, B.C.Y. & Haman, S.E.M.	439.67	479.67	73.0	379.0
8	Propane	440	38	Vaughan, W.E. & Collins, F.C.	491.67	815.67	14.7	664.3
	n-Pentane	540	80	Wichterle, I. & Vejrosta, J.	605.81	689.67	48.5	591.2
9	Isobutane	346	33	Besserer, G.J. & Robinson, D.B.	559.67	709.67	73.0	1042.0
	Carbon Dioxide	615	69	Weber, L. A.	559.67	709.67	72.1	1073.3

Table 3.1. Test Cases in Binary VLE Systems

Test Case #	Components	Data Set #	#Pt	Authors	T(min) (R)	T(max) (R)	P(min) (psia)	P(max) (psia)
1	Methane	447	57	Kohn, J.P. & Shim, J.	536.67	873.27	0.2	2845.2
	n-Hexane	461	104	Chen, R.J.J., Chappelear, P.S. & Kobayashi, R.	342.90	491.69	19.9	2650.0
2	Methane	441	135	Kobayashi, R. & Wichterle, I.	234.67	359.87	0.2	748.0
	Ethane	505	27	Kobayashi, R., Wichterle, I. & Chappelear, P.S.	344.16	354.98	646.9	728.3
		529	25	Miller, R.C., Hiza, M.J. & Kidnay, A.J.	288.00	324.00	3.1	476.9
3	Methane	385	15	Kirk, B.S. & Ziegler, W.T.	185.37	209.67	293.9	1322.6
	Hydrogen	570	25	Masuoka, H., Yorizane, M., Toyama, A. & Yoshimura, S.	185.67	293.67	147.0	2204.4
4	Methane	428	29	Lacey, W.N. & Sage, B.H.	529.67	709.67	40.0	1900.0
	n-Butane	545	80	Wang, R.H., Azarnoosh, A., McKetta, J.J. & Roberts, L.R.	379.67	739.67	26.0	1915.0
		501	21	Sage, B.H., Jacobs, J. & Wiese, H.C.	499.67	679.67	200.0	1700.0
		587	4	Sauer, R.N.	559.67	559.67	111.0	1803.0
		539	7	Thodos, G., Rigas, T.J. & Mason, D.F.	559.67	559.67	960.0	1861.0
5	Ethane	450	15	Hoshino, D., Nagahama, K., Hirata, M. & Konishi, H.	455.31	455.31	207.2	335.1
	Carbon Dioxide	566	17	Hakuta, T., Nagahama, K. & Suda, S.	491.67	491.67	348.3	577.6
		519	15	Kidnay, A.J., Phelps, R.E., Davalos, J. & Anderson, W.R.	450.00	450.00	188.8	309.6
6	Ethane	423	27	Robinson, D.B., Krishnan, T.R. & Kalra, H.	359.87	509.67	14.3	433.2
	Hydrogen Sulfide	796	45	Kalra, H. & Robinson, D.B.	359.87	509.67	9.5	442.7
7	Propane	341	15	Akers, W.W., Lipscomb, T.G. & Kelley, R.E.	419.67	491.67	15.0	507.0
	Carbon Dioxide	451	24	Nagahama, K., Hirata, M., Hoshino, D. & Konishi, H.	455.31	491.67	35.3	505.5
		512	21	Lu, B.C.Y. & Haman, S.E.M.	439.67	479.67	73.0	379.0
8	Propane	440	38	Vaughan, W.E. & Collins, F.C.	491.67	815.67	14.7	664.3
	n-Pentane	540	80	Wichterle, I. & Vejrosta, J.	605.81	689.67	48.5	591.2
9	Isobutane	346	33	Besserer, G.J. & Robinson, D.B.	559.67	709.67	73.0	1042.0
	Carbon Dioxide	615	69	Weber, L. A.	559.67	709.67	72.1	1073.3

Table 3.2. Test Cases in Ternary VLE System

Test Case #	Components	Data Set #	#Pt	Authors	T(min) (R)	T(max) (R)	P(min) (psia)	P(max) (psia)
1	Methane	641	196	Hong, J.H. & Kobayashi, R.	194.62	311.83	400.1	1500.4
	Carbon Monoxide	644	38	Herman & Kremer	216.00	252.00	419.2	725.2
	Hydrogen							

Step 3: Data Extraction

To export the data sets that are listed in Tables 3.1 and 3.2, an application, FRMEXPORT, was used. This application was created in Access 7.0 by Maase (13), and it was used to export mainly the VLE data sets in the GPA Database by giving just the data set numbers (DSN).

The exported files are in text format, and they are space delimited. A sample of these files can be found in Appendix B. These files can immediately be used as the input files of the Fortran program discussed in next chapter.

CHAPTER IV

EQUATION OF STATE VLE DATA MODEL

This chapter presents the basic equations of the SRK EOS and a general discussion of bubble-point calculations. These equations and calculations were used in the third methodology of the error analyses, which is called "data screening methodology". Also, the Fortran program is described.

The Equation-of-State Model

The SRK EOS was used in the data screening methodology (18):

$$P = \frac{RT}{v-b} - \frac{a}{v(v+b)} \tag{3.1}$$

where the constant b is (14)

$$b = \left(\sum x_i b_i\right) \tag{3.2}$$

and

$$b_i = 0.08664 \frac{RT_{c,i}}{P_{c,i}} \tag{3.3}$$

The parameter a depends on temperature according to (19)

$$a = \left(\sum x_i a_i^{0.5}\right)^2 \tag{3.4}$$

where

$$a_i = a(T_{c,i})\alpha_i(T) \tag{3.5}$$

and

$$a(T_{c,i}) = 0.42748\frac{R^2T_{c,i}^2}{P_{c,i}} \tag{3.6}$$

For normal fluids (19),

$$\alpha_i(T) = \left[1 + \left(0.480 + 1.574\omega_i - 0.176\omega_i^2\right)\left(1 - \sqrt{T/T_{c,i}}\right)\right]^2 \tag{3.7}$$

where ω is the accentric factor.

Equation (3.1) can also be written as (18):

$$Z^3 - Z^2 + Z\left(A - B - B^2\right) - AB = 0 \tag{3.8}$$

where

$$Z = \frac{PV}{RT}, \tag{3.9}$$

$$A = \frac{aP}{R^2T^2}, \tag{3.10}$$

and

$$B = \frac{bP}{RT} \tag{3.11}$$

Equation (3.2) and (3.4) were the original generalized mixing rules Soave (18) used. These mixing rules could be applied with acceptable results to mixtures of non-polar fluids, such as hydrocarbons, nitrogen, and carbon monoxide.

However, some empirical corrections are necessary for systems containing carbon dioxide, hydrogen sulfide and polar compounds, for which large deviations were obtained, although the vapor pressures of the single pure components were reproduced well (18). Therefore the following modifications in the mixing rules are employed (18):

$$a = (1 - k_{ij})(a_i a_j)^{0.5} \tag{3.12}$$

where k_{ij} is a binary interaction parameter to be determined empirically and characterizes the differences in molecular size of component i and j (18).

Bubble-Point Calculations

The fugacity coefficient of the liquid phase, $\hat{\phi}_L$, and vapor phase, $\hat{\phi}_v$, can be calculated by using the following equation (7):

$$\ln \hat{\phi}_i = -\ln(Z - B) + (Z - 1)B_i' - \frac{A}{B}\left[A_i' - B_i'\right]\ln\left(1 + \frac{B}{Z}\right) \tag{3.13}$$

where

$$B_i' = \frac{b_i}{b} \quad A_i' = \frac{1}{a}\left[2a_i \sum_j^N x_j a_j^{0.5}(1 - k_{ij})\right] \tag{3.14}$$

The calculated values of $\hat{\phi}_i^l$ and $\hat{\phi}_i^v$ can be used in dew-point, bubble-point, and flash calculations. Since the feed composition values are not available for the majority of data points of VLE systems, flash calculations were not selected. On the other hand, bubble-point and dew-point calculations are equally accurate in the model. These two types of calculations can be further classified as pressure dependent and temperature dependent calculations. In the data screening methodology, pressure dependent bubble-point calculations were used to predict the system pressures, P, and vapor-phase compositions, y_i.

Oklahoma State University Library

These two values were predicted by specifying values of temperature, T; liquid-phase compositions, x_i; and the physical-property data necessary for evaluation of all EOS parameters. Pressures can be computed by using Equation (3.1) whereas y-values were calculated from the following equation:

$$y_i = K_i x_i \tag{3.15}$$

where the K-value, K_i, is given by

$$K_i = \frac{\hat{\phi}_i^{l}}{\hat{\phi}_i^{v}} \tag{3.16}$$

One goal of calculating these values in the data screening methodology is to check the reliability of F-tests in the model-predictions-check methodology. Also, data discrepancies and data meriting further examination were identified. These analyses are described as point-by-point analyses.

In these analyses, certain criteria were used to identify the data records that need further examination. The data records noted were:

(i) Data points showing deviations in calculated pressure and vapor composition values that are greater than twice the root-mean-squared error (RMSE) for the entire data set.

(ii) Data values exhibiting gross systematic errors; these may be identified by the trends in deviations of the data sets.

(iii) Data points showing an abrupt change in the deviation sign.

These criteria were developed by Rastogi (21). They were used to identify data records showing deviations between the reported and predicted values that were larger

Oklahoma State University Library

than expected. The results of the data screening methodology can be found in the following chapter.

Fortran Program Used

A modified version of the GPA*Sim program (5) was used in this study. This code was modified to include the data screening methodology. Deviations between the model predictions and experimental data were calculated for pressures and vapor compositions.

Also, twice the RMSE in percent were calculated to screen the pressure data records. The range of the pressure measurements is significant, so RMSE in percent is recommended. On the other hand, twice the RMSE were calculated to screen the vapor composition data records, since the range of this variable is not significant.

CHAPTER V

RESULTS AND DISCUSSIONS OF VLE DATA EVALUATION

The evaluation of the VLE data was performed using three methodologies of the error analyses used in this study. They are population-check methodology, model-predictions-check methodology, and data screening methodology. The main goals of the first two methodologies based on F-tests were discussed in Chapter II. For the data screening methodology, the main goal was discussed in Chapter IV. About 1,800 data points were tested with these three methodologies. Typical test cases are also discussed in this chapter.

Population-Check Methodology

Two binary systems were selected as examples for the application of the population-check methodology. The first is the methane-ethane binary system and consists of three data sets. The second binary system contains methane and n-butane with five data sets. All of the data sets have different investigators. Refer to Table 3.1 in Chapter III for more details.

Tables 5.1 and 5.2 on the following pages present the results of the population checks on these two binary systems. The null hypothesis in the tables, H_0, states that the variances of the samples are equal to each other.

Methane-Ethane Binary System

This system includes Data Sets 441, 505 and 529. They were tested in three pairs. The results presented in Table 5.1 are based on one-sided tests at significant levels of 0.05 and 0.01.

The results clearly show that the calculated F-ratios of P, T, x, and y variables from Data Sets 505 and 441 are much larger than the tabulated F-values. Therefore, sample variances of Data Set 505 are not equal to the sample variances of Data Set 441 and major differences are expected in the data measurements of these two data sets.

The same key points can be applied to Data Sets 505 and 529. Their variances are largely different from each other as shown in Table 5.1.

A different case was found by comparing the calculated F-ratios from Data Set 529 and 441 to the tabulated F-values. The results indicate that the variances of x and y variables from these two data sets are equal to each other. Thus, they are considered to come from the same population, and major differences are not expected in these two variables. For T and P variables, the hypothesis is rejected both at 0.05 and 0.01 significant levels under the one-sided tests. Thus, they did not come from the same population, and major differences are expected in these two variables.

Table 5.1 F-Tests for the Population-Check Methodology Applied to the Methane-Ethane Binary VLE System

Data Set #	Comp#1	Comp#2	Tests	T	P	x_1	x_2	y_1	y_2
441	Methane	Ethane	$F_{calculated}$	89.66	130.33	384.52	384.52	692.76	693.31
505			$F_{0.05}$	1.75	1.75	1.75	1.75	1.75	1.75
			H_o	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected
			$F_{0.01}$	2.23	2.23	2.23	2.23	2.23	2.23
			H_o	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected
529	Methane	Ethane	$F_{calculated}$	3.51	4.12	1.24	1.24	1.10	1.10
441			$F_{0.05}$	1.79	1.79	1.79	1.79	1.59	1.59
			H_o	Rejected	Rejected	Accepted	Accepted	Accepted	Accepted
			$F_{0.01}$	2.31	2.31	2.31	2.31	1.92	1.92
			H_o	Rejected	Rejected	Accepted	Accepted	Accepted	Accepted
505	Methane	Ethane	$F_{calculated}$	25.58	31.60	309.92	309.92	762.99	762.99
529			$F_{0.05}$	1.95	1.95	1.95	1.95	1.96	1.96
			H_o	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected
			$F_{0.01}$	2.58	2.58	2.58	2.58	2.60	2.60
			H_o	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected

Table 5.2 F-Tests for the Population-Check Methodology Applied to the Methane-n-Butane Binary VLE System

Data Set #	Comp#1	Comp#2	Tests	T	P	x_1	x_2	y_1	y_2
428	Methane	n-Butane	$F_{calculated}$	3.27	1.18	1.37	1.37	1.04	1.04
545			$F_{0.05}$	1.75	1.62	1.75	1.75	1.62	1.62
			H_o	Rejected	Accepted	Accepted	Accepted	Accepted	Accepted
			$F_{0.01}$	2.22	2.01	2.22	2.22	2.01	2.01
			H_o	Rejected	Accepted	Accepted	Accepted	Accepted	Accepted
501	Methane	n-Butane	$F_{calculated}$	2.93	1.09	2.06	1.84	1.73	1.66
545			$F_{0.05}$	1.92	1.92	1.92	1.92	1.92	1.92
			H_o	Rejected	Accepted	Rejected	Accepted	Accepted	Accepted
			$F_{0.01}$	2.56	2.56	2.56	2.56	2.56	2.56
			H_o	Rejected	Accepted	Accepted	Accepted	Accepted	Accepted
428	Methane	n-Butane	$F_{calculated}$	1.11	1.29	1.35	1.35	1.72	1.72
501			$F_{0.05}$	1.96	2.06	2.06	2.06	2.06	2.06
			H_o	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
			$F_{0.01}$	2.60	2.80	2.80	2.80	2.80	2.80
			H_o	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted

Note: NV means F-ratio cannot be found in this variable because one of the data sets is isothermal and so the variance is equal to zero.

Table 5.2 F-Tests for the Population-Check Methodology Applied to the Methane-n-Butane Binary VLE System (continued)

Data Set #	Comp#1	Comp#2	Tests	T	P	x_1	x_2	y_1	y_2
539	Methane	n-Butane	$F_{calculated}$	NV	2.57	3.36	3.36	39.67	39.67
545			$F_{0.05}$		3.72	3.72	3.72	3.72	3.72
			H_o		Accepted	Accepted	Accepted	Rejected	Rejected
			$F_{0.01}$		7.02	7.02	7.02	7.02	7.02
			H_o		Accepted	Accepted	Accepted	Rejected	Rejected
587	Methane	n-Butane	$F_{calculated}$	NV	1.82	1.00	1.00	1.59	1.59
545			$F_{0.05}$		2.72	2.72	2.72	8.57	8.57
			H_o		Accepted	Accepted	Accepted	Accepted	Accepted
			$F_{0.01}$		4.04	4.04	4.04	26.3	26.3
			H_o		Accepted	Accepted	Accepted	Accepted	Accepted

Note: NV means F-ratio cannot be found in this variable because one of the data sets is isothermal and so the variance is equal to zero.

To further investigate the variables that did not come from the same population, temperature and pressure ranges of these three data sets are the main factors that caused the sample variances significantly differ from each other.

Methane-n-Butane Binary System

Five data sets were tested in this system. They are Data Sets 428, 501, 539, 545, and 587. These data sets were tested in five pairs. The results of one-sided tests are given in Table 5.2.

A similar illustration of one-sided tests can be adopted to these five tested groups which involved five data sets. Based on the results, most of the variables in these test cases are expected to come from the same populations. For the test case that includes Data Sets 428 and 545, the only variable that did not come from the same population is T.

On the other hand, T and x_1 variables did not come from the same population in the test case that included Data Sets 501 and 545. For the case that included Data Sets 539 and 545, the only variable that did not come from the same population is y. Therefore, major differences are expected only in the variables that did not come from the same population in these five data sets. To further investigate the major differences of these variables, different investigated temperature range of the data sets in this system is considered as one of the main factors.

Discussion

The analyses presented above for the two binary systems highlight the key points of the population-check methodology. Based on the outcomes of one-sided tests, variables that came from the same population in test cases were identified.

As mentioned in Chapter II, this screening tool provides a quick way to determine whether there are any major differences in the data measurements of the same-component data sets reported by different investigators. This can be done without checking the individual data points of the data sets. However, the results of F-tests in this methodology need to be further analyzed by the model-predictions methodology.

A similar analysis was employed to screen another 18 binary systems and 1 ternary system. Table 5.3 presents the outcomes of F-tests in binary systems. Thirty-four data sets were analyzed. The outcomes of F-tests in the ternary system are presented in Table 5.4. Two data sets were analyzed. Detailed calculations and comparisons of these variance tests can be found in the tables of Appendix C.

These results indicated that most of the variables in same-component data sets of more than 65% of the VLE systems are expected to come from the same populations. Thus, major differences are not expected in the measurements of those variables. For the sample variables that did not come from the same populations, different investigated temperature and pressure ranges are the main factors for the major differences.

Table 5.3. F-Tests for the Population-Check Methodology Applied to Binary VLE Systems

Case #	Data Set #	Comp#1	Comp#2	Probability Levels	T	P	x_1	x_2	y_1	y_2
1	447	Methane	n-Hexane	0.05	-	-	-	-	-	-
	461			0.01	-	+	-	-	-	-
2	385	Methane	Hydrogen	0.05	-	-	-	-	-	-
	570			0.01	-	+	-	-	-	-
3a	450	Ethane	Carbon Dioxide	0.05	+	-	+	+	+	+
	566			0.01	+	+	+	+	+	+
3b	519	Ethane	Carbon Dioxide	0.05	NV	-	+	+	+	+
	566			0.01		-	+	+	+	+
4	423	Ethane	Hydrogen Sulfide	0.05	+	+	+	+	+	+
	796			0.01	+	+	+	+	+	+
5a	341	Propane	Carbon Dioxide	0.05	-	+	+	+	+	+
	451			0.01	+	+	+	+	+	+
5b	341	Propane	Carbon Dioxide	0.05	+	-	-	-	-	-
	512			0.01	+	+	+	+	-	-

Note: NV means F-ratio cannot be found on this variable because one of the data sets is isothermal and so the variance is equal to zero.

The + signs indicate the variables are expected to come from the same population whereas – signs indicate the variables are not expected to come from the same population.

Table 5.3. F-Tests for the Population-Check Methodology Applied in Binary VLE Systems (continued)

Case #	Data Set #	Comp#1	Comp#2	Probability Levels	T	P	x_1	x_2	y_1	y_2
5c	451	Propane	Carbon Dioxide	0.05	+	-	-	-	-	-
	512			0.01	+	+	+	+	-	-
6	440	Propane	n-Pentane	0.05	-	-	+	+	-	-
	540			0.01	-	-	+	+	-	-
7	346	Isobutane	Carbon Dioxide	0.05	+	+	+	+	+	+
	615			0.01	+	+	+	+	+	+
8	350	Isopentane	Cyclohexane	0.05	+	NV	+	+	+	+
	790			0.01	+		+	+	+	+
9	351	Isopentane	Methylcyclohexane	0.05	-	NV	+	+	+	+
	791			0.01	+		+	+	+	+
10	792	n-Heptane	Cyclohexane	0.05	+	NV	+	+	+	+
	828			0.01	+		+	+	+	+
11	433	n-Heptane	Toluene	0.05	+	NV	+	+	+	+
	823			0.01	+		+	+	+	+

Note: NV means F-ratio cannot be found on this variable because one of the data sets is isobaric and so the variance is equal to zero.

The + signs indicate the variables are expected to come from the same population whereas – signs indicate the variables are not expected to come from the same population.

Table 5.3. F-Tests for the Population-Check Methodology Applied in Binary VLE Systems (continued)

Case #	Data Set #	Comp#1	Comp#2	Probability Levels	T	P	x_1	x_2	y_1	y_2
12	507	Carbon Monoxide	Carbon Dioxide	0.05	+	+	+	+	+	+
	557			0.01	+	+	+	+	+	+
13	500	Carbon Monoxide	Hydrogen	0.05	+	-	-	-	+	+
	571			0.01	+	-	+	+	+	+
14	349	Carbon Dioxide	Hydrogen Sulfide	0.05	-	-	+	+	+	+
	432			0.01	-	+	+	+	+	+
15	522	Carbon Dioxide	Nitrogen	0.05	NV	+	+	+	+	+
	569			0.01		+	+	+	+	+
16	546	Hydrogen	Nitrogen	0.05	-	-	-	-	+	+
	573			0.01	-	-	-	-	+	+

Note: NV means F-ratio cannot be found on this variable because one of the data sets is isothermal and so the variance is equal to zero.

The + signs indicate the variables are expected to come from the same population whereas – signs indicate the variables are not expected to come from the same population.

Table 5.4. F-Tests for the Population-Check Methodology Applied to a Ternary VLE System

Case #	Data Set #	Comp#1	Comp#2	Comp#3	Probability Levels	T	P	x_1	x_2	x_3	y_1	y_2	y_3
1	641	Hydrogen	Methane	Carbon	0.05	-	-	-	+	+	+	-	+
	644			Monoxide	0.01	-	-	-	+	+	+	-	+

Note: The + signs indicate the variables are expected to come from the same population whereas – signs indicate the variables are not expected to come from the same population.

Model-Predictions-Check Methodology

Two binary systems were selected for the discussion of this section. They are methane-ethane and methane-n-butane binary systems. Data Sets 441, 505 and 529 are for the methane-ethane binary system; Data Sets 545, 428, 501, 587, and 539 are for the methane-n-butane binary system. Table 5.5 and 5.6 present the results of these systems.

Methane-Ethane Binary System

For all the tested data sets in methane-ethane binary system, the calculated F-ratios for the P and y variables exhibit large differences compared to the tabulated F-values at 0.05 and 0.01 significant levels. This clearly indicates that the explained variances for P and y variables are significant when compared to the unexplained variances. Therefore, the model predictions in these data sets are reasonable.

As a closer look at the F-tests of population-check methodology, the model predictions are considered reasonable regardless of the major differences between Data Sets 441 and 505 and between Data Sets 505 and 529. Therefore, the SRK EOS did predict well not only in the data sets that came from the same populations but also data sets that did not come from the same populations.

Methane-n-Butane Binary System

In this system, the explained variances for P and y variables in all the five data sets are significant at 0.05 and 0.01 significant levels, when compared to the unexplained variances. Thus, the model predictions are reasonable in these data sets.

Table 5.5. F-Tests for the Model-Predictions-Check Methodology Applied to Methane-Ethane Binary System

DSN	Comp#1	Comp#2	Tests	P	y_1	y_2
441	Methane	Ethane	$F_{calculated}$	488315	44802	44802
			$F_{0.05}$	3.92	3.92	3.92
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	6.84	6.84	6.84
			H_o	Accepted	Accepted	Accepted
505	Methane	Ethane	$F_{calculated}$	540	76	76
			$F_{0.05}$	4.24	4.24	4.24
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.77	7.77	7.77
			H_o	Accepted	Accepted	Accepted
529	Methane	Ethane	$F_{calculated}$	20509	15095	15095
			$F_{0.05}$	4.28	4.28	4.28
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.88	7.88	7.88
			H_o	Accepted	Accepted	Accepted

Oklahoma State University Library

Table 5.6. **F-Tests for the Model-Predictions-Check Methodology Applied to Methane-n-Butane Binary System**

DSN	Comp#1	Comp#2	Tests	P	y_1	y_2
545	Methane	n-Butane	$F_{calculated}$	2752	1053	1053
			$F_{0.05}$	3.97	3.97	3.97
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.00	7.00	7.00
			H_o	Accepted	Accepted	Accepted
428	Methane	n-Butane	$F_{calculated}$	1722	905	905
			$F_{0.05}$	4.21	4.21	4.21
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.68	7.68	7.68
			H_o	Accepted	Accepted	Accepted
501	Methane	n-Butane	$F_{calculated}$	240	294	294
			$F_{0.05}$	4.38	4.38	4.38
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	8.18	8.18	8.18
			H_o	Accepted	Accepted	Accepted
587	Methane	n-Butane	$F_{calculated}$	281	99	99
			$F_{0.05}$	18.51	18.51	18.51
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	98.49	98.49	98.49
			H_o	Accepted	Accepted	Accepted
539	Methane	n-Butane	$F_{calculated}$	249	19	19
			$F_{0.05}$	6.61	6.61	6.61
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	16.26	16.26	16.26
			H_o	Accepted	Accepted	Accepted

Based on the F-tests in population-check methodology, major differences are expected in the variable y, when Data Set 539 is compared to the Data Set 545 in this system. However, the model predictions of this variable are reasonable in both of these data sets regardless of the major differences. The SRK EOS also predicted equally well in those P and y variables that came from the same populations in this system.

Discussion

The analyses presented above for the two binary systems highlight the key points of the model-predictions-check methodology. Based on the results, one can further investigate F-tests in the population-check methodology and determine whether the SRK EOS predicted equally well in those variables that came from the same populations. Also, one can determine whether the SRK EOS is suitable to model the VLE data sets in the GPA Database.

A similar analysis was applied to another 39 data sets in binary systems and 2 data sets in a ternary system. Tables 5.7 and 5.8 present the calculated F-ratios and comparisons of F-tests in these data sets.

The results revealed that the model predictions are reasonable in almost all the selected VLE data sets. Only model predictions of P variable in three data sets are not reasonable. They are Data Sets 350, 351 and 569. In these data sets, the values of the unexplained variances, $S^2(\hat{P})$, are 12, 19 and 141358 respectively. These numbers are considerably high, when compared to the explained variances. These results were further analyzed in the data screening methodology.

Table 5.7. F-Tests for the Model-Predictions-Check Methodology Applied to Binary VLE Systems

DSN	Comp#1	Comp#2	Tests	P	y_1	y_2
447	Methane	n-Hexane	$F_{calculated}$	516	43	43
			$F_{0.05}$	4.02	4.02	4.02
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.12	7.12	7.12
			H_o	Accepted	Accepted	Accepted
461	Methane	n-Hexane	$F_{calculated}$	2426	517	517
			$F_{0.05}$	3.94	3.94	3.94
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	6.90	6.90	6.90
			H_o	Accepted	Accepted	Accepted
570	Methane	Hydrogen	$F_{calculated}$	77	229	229
			$F_{0.05}$	4.28	4.28	4.28
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.88	7.88	7.88
			H_o	Accepted	Accepted	Accepted
385	Methane	Hydrogen	$F_{calculated}$	8	59	59
			$F_{0.05}$	4.67	4.67	4.67
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	9.07	9.07	9.07
			H_o	Rejected	Accepted	Accepted
450	Ethane	Carbon Dioxide	$F_{calculated}$	109	6099	6099
			$F_{0.05}$	4.67	4.67	4.67
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	9.07	9.07	9.07
			H_o	Accepted	Accepted	Accepted

Table 5.7. F-Tests for the Model-Predictions-Check Methodology Applied to Binary VLE Systems (continued)

DSN	Comp#1	Comp#2	Tests	P	y_1	y_2
566	Ethane	Carbon Dioxide	$F_{calculated}$	546	16509	16509
			$F_{0.05}$	4.54	4.54	4.54
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	8.68	8.68	8.68
			H_o	Accepted	Accepted	Accepted
519	Ethane	Carbon Dioxide	$F_{calculated}$	154	5020	5020
			$F_{0.05}$	4.67	4.67	4.67
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	9.07	9.07	9.07
			H_o	Accepted	Accepted	Accepted
423	Ethane	Hydrogen Sulfide	$F_{calculated}$	6887	850	850
			$F_{0.05}$	4.24	4.24	4.24
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.77	7.77	7.77
			H_o	Accepted	Accepted	Accepted
796	Ethane	Hydrogen Sulfide	$F_{calculated}$	18645	2781	2781
			$F_{0.05}$	4.07	4.07	4.07
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.27	7.27	7.27
			H_o	Accepted	Accepted	Accepted
341	Propane	Carbon Dioxide	$F_{calculated}$	60	103	103
			$F_{0.05}$	4.67	4.67	4.67
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	9.07	9.07	9.07
			H_o	Accepted	Accepted	Accepted

Oklahoma State University Library

Table 5.7. F-Tests for the Model-Predictions-Check Methodology Applied to Binary VLE Systems (continued)

DSN	Comp#1	Comp#2	Tests	P	y_1	y_2
451	Propane	Carbon Dioxide	$F_{calculated}$	3005	2190	2190
			$F_{0.05}$	4.30	4.30	4.30
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.94	7.94	7.94
			H_o	Accepted	Accepted	Accepted
512	Propane	Carbon Dioxide	$F_{calculated}$	1558	2212	2212
			$F_{0.05}$	4.38	4.38	4.38
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	8.18	8.18	8.18
			H_o	Accepted	Accepted	Accepted
440	Propane	n-Pentane	$F_{calculated}$	1306	1541	1541
			$F_{0.05}$	4.11	4.11	4.11
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.39	7.39	7.39
			H_o	Accepted	Accepted	Accepted
540	Propane	n-Pentane	$F_{calculated}$	46820	87917	87917
			$F_{0.05}$	3.97	3.97	3.97
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.00	7.00	7.00
			H_o	Accepted	Accepted	Accepted
346	Isobutane	Carbon Dioxide	$F_{calculated}$	3654	5638	5638
			$F_{0.05}$	4.16	4.16	4.16
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.53	7.53	7.53
			H_o	Accepted	Accepted	Accepted

Oklahoma State University Library

Table 5.7. F-Tests for the Model-Predictions-Check Methodology Applied to Binary VLE Systems (continued)

DSN	Comp#1	Comp#2	Tests	P	y_1	y_2
615	Isobutane	Carbon Dioxide	$F_{calculated}$	1762	24194	24194
			$F_{0.05}$	3.99	3.99	3.99
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.04	7.04	7.04
			H_o	Accepted	Accepted	Accepted
350	Isopentane	Cyclohexane	$F_{calculated}$	5	245	245
			$F_{0.05}$	6.61	6.61	6.61
			H_o	Rejected	Accepted	Accepted
			$F_{0.01}$	16.26	16.26	16.26
			H_o	Rejected	Accepted	Accepted
790	Isopentane	Cyclohexane	$F_{calculated}$	26	1179	1179
			$F_{0.05}$	4.20	4.20	4.20
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.64	7.64	7.64
			H_o	Accepted	Accepted	Accepted
351	Isopentane	Methylcyclohexane	$F_{calculated}$	5	3962	3962
			$F_{0.05}$	6.61	6.61	6.61
			H_o	Rejected	Accepted	Accepted
			$F_{0.01}$	16.26	16.26	16.26
			H_o	Rejected	Accepted	Accepted
791	Isopentane	Methylcyclohexane	$F_{calculated}$	45	1109	1109
			$F_{0.05}$	4.05	4.05	4.05
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.21	7.21	7.21
			H_o	Accepted	Accepted	Accepted

Oklahoma State University Library

Table 5.7. F-Tests for the Model-Predictions-Check Methodology Applied to Binary VLE Systems (continued)

DSN	Comp#1	Comp#2	Tests	P	y_1	y_2
792	n-Heptane	Cyclohexane	$F_{calculated}$	23	4570	4570
			$F_{0.05}$	4.24	4.24	4.24
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.77	7.77	7.77
			H_o	Accepted	Accepted	Accepted
828	n-Heptane	Cyclohexane	$F_{calculated}$	15	9911	9911
			$F_{0.05}$	4.45	4.45	4.45
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	8.40	8.40	8.40
			H_o	Accepted	Accepted	Accepted
433	n-Heptane	Toluene	$F_{calculated}$	17	37789	37789
			$F_{0.05}$	4.38	4.38	4.38
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	8.18	8.18	8.18
			H_o	Accepted	Accepted	Accepted
823	n-Heptane	Toluene	$F_{calculated}$	20	39623	39623
			$F_{0.05}$	4.30	4.30	4.30
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.94	7.94	7.94
			H_o	Accepted	Accepted	Accepted
507	Carbon Monoxide	Carbon Dioxide	$F_{calculated}$	754	2426	2426
			$F_{0.05}$	4.35	4.35	4.35
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	8.10	8.10	8.10
			H_o	Accepted	Accepted	Accepted

Oklahoma State University Library

Table 5.7. F-Tests for the Model-Predictions-Check Methodology Applied to Binary VLE Systems (continued)

DSN	Comp#1	Comp#2	Tests	P	y_1	y_2
557	Carbon Monoxide	Carbon Dioxide	$F_{calculated}$	1057	1605	1605
			$F_{0.05}$	4.11	4.11	4.11
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.39	7.39	7.39
			H_o	Accepted	Accepted	Accepted
500	Carbon Monoxide	Hydrogen	$F_{calculated}$	129	3169	3169
			$F_{0.05}$	3.92	3.92	3.92
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	6.84	6.84	6.84
			H_o	Accepted	Accepted	Accepted
571	Carbon Monoxide	Hydrogen	$F_{calculated}$	8	782	782
			$F_{0.05}$	4.30	4.30	4.30
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.94	7.94	7.94
			H_o	Accepted	Accepted	Accepted
349	Carbon Dioxide	Hydrogen Sulfide	$F_{calculated}$	6633	29008	29008
			$F_{0.05}$	3.96	3.96	3.96
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	6.96	6.96	6.96
			H_o	Accepted	Accepted	Accepted
432	Carbon Dioxide	Hydrogen Sulfide	$F_{calculated}$	14993	15650	15650
			$F_{0.05}$	3.96	3.96	3.96
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	6.96	6.96	6.96
			H_o	Accepted	Accepted	Accepted

Oklahoma State University Library

Table 5.7. F-Tests for the Model-Predictions-Check Methodology Applied to Binary VLE Systems (continued)

DSN	Comp#1	Comp#2	Tests	P	y_1	y_2
522	Carbon Dioxide	Nitrogen	$F_{calculated}$	5	91	91
			$F_{0.05}$	4.13	4.13	4.13
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.44	7.44	7.44
			H_o	Rejected	Accepted	Accepted
569	Carbon Dioxide	Nitrogen	$F_{calculated}$	3	39	39
			$F_{0.05}$	4.60	4.60	4.60
			H_o	Rejected	Accepted	Accepted
			$F_{0.01}$	8.86	8.86	8.86
			H_o	Rejected	Accepted	Accepted
546	Hydrogen	Nitrogen	$F_{calculated}$	35	76	76
			$F_{0.05}$	4.54	4.54	4.54
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	8.68	8.68	8.68
			H_o	Accepted	Accepted	Accepted
573	Hydrogen	Nitrogen	$F_{calculated}$	23	33	33
			$F_{0.05}$	4.17	4.17	4.17
			H_o	Accepted	Accepted	Accepted
			$F_{0.01}$	7.56	7.56	7.56
			H_o	Accepted	Accepted	Accepted

Oklahoma State University Library

Table 5.8. F-Tests for the Model-Predictions-Check Methodology Applied to a Ternary VLE System

DSN	Comp#1	Comp#2	Comp#3	Tests	P	y_1	y_2	y_3
641	Hydrogen	Methane	Carbon Monoxide	$F_{calculated}$	224	8774	17511	7567
				$F_{0.05}$	3.90	3.90	3.90	3.90
				H_o	Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	6.80	6.80	6.80	6.80
				H_o	Accepted	Accepted	Accepted	Accepted
644	Hydrogen	Methane	Carbon Monoxide	$F_{calculated}$	16	1288	1487	2671
				$F_{0.05}$	4.11	4.11	4.11	4.11
				H_o	Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	7.39	7.39	7.39	7.39
				H_o	Accepted	Accepted	Accepted	Accepted

Oklahoma State University Library

For Data Sets 350 and 351, F-ratios cannot be found in the population-check methodology due to the constant pressure condition. For Data Set 569, the SRK EOS did not predict equally well in the P variable, even though there are no major differences between Data Sets 569 and 522.

As an overview, the results of the model-predictions methodology revealed that the SRK EOS did predict well in the data sets that came from the same populations and also data sets that did not came from the same population. To test the reliability of F-tests in this methodology, the results of data screening methodology are discussed in the following sections.

Data Screening Methodology

The following sections focus on discussing the results of the data screening methodology. Typical test case systems were selected and discussed to demonstrate the checking of the reliability of F-tests in the model-predictions-check methodology. Also, data discrepancies and data meriting further examination were identified.

Two binary systems that involved eight data sets were selected for the discussion. They are methane-ethane and methane-n-butane binary systems. Tables 5.9 and 5.10 give a summary of the data records that have been flagged in the data sets of these two systems.

Table 5.9. Summary of Possible Outliers in the Methane-Ethane Binary System

DSN	Comp#1	Comp#2	Raw/Smooth	Total Pts	# $P_{flagged}$	# $y(_1)_{flagged}$	# $y(_2)_{flagged}$
441	Methane	Ethane	R	135	7	11	11
505	Methane	Ethane	R	27	0	0	0
529	Methane	Ethane	R	25	2	3	3

Table 5.10. Summary of Possible Outliers in the Methane-n-Butane Binary System

DSN	Comp#1	Comp#2	Raw/Smooth	Total Pts	# $P_{flagged}$	# $y(_1)_{flagged}$	# $y(_2)_{flagged}$
545	Methane	n-Butane	R	78	4	2	2
428	Methane	n-Butane	Unknown	29	0	0	0
501	Methane	n-Butane	R	21	1	1	1
587	Methane	n-Butane	Unknown	4	0	0	0
539	Methane	n-Butane	R	7	0	0	0

Methane-Ethane Binary System

Data Sets 441, 505 and 529 were chosen in this system. Among three of them, some of the data records in Data Sets 441 and 529 were flagged based on the criteria discussed in Chapter IV.

In Data Set 441, 135 data points were tested and evaluated. All of these data points are raw data. Data records of pressure and vapor compositions that require further examination in this system are presented in Tables 5.11 and 5.12.

For the pressure predictions, most of the pressure records at different isotherms show a consistently low-percentage deviation, except for seven pressure records shown in Table 5.11. They have been flagged for showing deviations exceeding twice the RMSE value for the entire data set.

At 234 R, the pressure record, 28 psia, also was marked for showing abrupt change in deviation sign compared to neighboring records on this isotherm. Three pressure records at 284 R have also been flagged for showing gross systematic deviations.

For the vapor composition predictions, eleven data points in this system were identified as data exhibiting "higher-than-expected" deviations. They are presented in Table 5.12. All have been flagged for showing deviations exceeding twice the RMSE value for the entire data set.

Since most of the data records at different isotherms show a consistently low deviation, the model predictions can be inferred as reasonable, which supports the expectation made in the model-predictions methodology. For more evidence, the

Table 5.11. Pressure Data Records that Have Been Flagged in Data Set 441

Data Point #	Temp. (R)	Press. (psia)	Criteria for Outliers
2	234.67	28.0	(i) , (iii)
17	284.67	25.8	(i) , (ii)
18	284.67	28.8	(i) , (ii)
20	284.67	50.0	(i) , (ii)
56	341.37	35.8	(i)
93	346.27	100.0	(i)
122	359.87	100.0	(i)

Table 5.12. Vapor Composition Data Records that Have Been Flagged in Data Set 441

Data Point #	Temp. (R)	Y_1	Criteria for Outliers	Y_2	Criteria for Outliers
29	309.67	0.7681	(i)	0.2319	(i)
40	334.97	0.5585	(i)	0.4415	(i)
56	341.37	0.4600	(i)	0.5400	(i)
57	341.37	0.6741	(i)	0.3259	(i)
71	343.57	0.4724	(i)	0.5276	(i)
72	344.57	0.6513	(i)	0.3487	(i)
91	346.27	0.4339	(i)	0.5661	(i)
92	347.27	0.6358	(i)	0.3642	(i)
120	359.87	0.3005	(i)	0.6995	(i)
121	359.87	0.5098	(i)	0.4902	(i)
122	359.87	0.6800	(i)	0.3200	(i)

deviation signs are consistent compared to the neighboring records on each isotherm of this data set, and SRK EOS is well known to predict reasonably in hydrocarbon systems.

In Data Set 529, two pressure data records have been flagged at 288 R and a sudden increase in percentage deviations of these two data records was observed. Three vapor composition data records have also been flagged at 288 R due to the consistently high deviations. All of these data records have been flagged for showing deviations exceeding twice the RMSE value for the entire data set as shown in Table 5.13 and 5.14.

Despite the data records that have been flagged for showing high deviations, the SRK EOS predicted well in this data set that has moderate pressure conditions. Also, a consistently low deviation was observed in other data records. Therefore, F-tests in model-predictions methodology are considered reliable for this data set.

For Data Set 505, there are no data records requiring further examination based on the criteria used in this methodology. A consistently low deviation was observed and the SRK EOS is considered applicable to this data set that has moderate pressure condition. Thus, F-tests in model-predictions methodology are also considered reliable for this data set.

Methane-n-Butane Binary System

In this system, Data Sets 545, 428, 501, 587 and 539 were tested and evaluated. No data records were flagged based on the criteria for Data Sets 428, 587 and 539 as shown in Table 5.10.

Table 5.13. Pressure Data Records that Have Been Flagged in Data Set 529

Data Point #	Temp. (R)	Press. (psia)	Criteria for Outliers
2	288.00	29.4	(i)
3	288.00	29.5	(i)

Table 5.14. Vapor Composition Data Records that Have Been Flagged in Data Set 529

Data Point #	Temp. (R)	Y_1	Criteria for Outliers	Y_2	Criteria for Outliers
1	288.00	0.8270	(i)	0.1730	(i)
2	288.00	0.8970	(i)	0.1030	(i)
3	288.00	0.8986	(i)	0.1014	(i)

The model predictions in Data Sets 428, 587 and 539 are considered reasonable, since low deviations were observed in these data sets. These results also show that the F-tests in model-predictions-check methodology are reliable.

For Data Set 545, data records of pressure and vapor compositions that require further examination are presented in Tables 5.15 and 5.16. Based on the results of Table 5.15, four pressure data records have been flagged at 500 R and 380 R. In Table 5.16, two vapor composition data records have been flagged at 500 R and 440 R. These data records of pressure and vapor compositions have been flagged for showing deviations exceeding twice the RMSE value for the entire data set.

The model predictions in this data set are considered reasonable, since deviations in the data records of the isotherms are consistently low. Also, there is no abrupt change of deviation signs.

In Data Set 501, only one pressure data record and one vapor composition data record have been flagged at 560 R. They have been flagged for showing deviations exceeding twice the RMSE value for the entire data set as shown in Tables 5.17 and 5.18.

Although the pressure condition of this data set is quite high, the model predictions are considered reasonable due to the consistently low deviations. Also, there is no abrupt change of sign deviation, except for the flagged data at 560 R.

The results for these two binary systems indicate that the F-tests in the model-predictions-check methodology are reliable in this study.

Table 5.15. Pressure Data Records that Have Been Flagged in Data Set 545

Data Point #	Temp. (R)	Press. (psia)	Criteria for Outliers
11	499.67	53.0	(i)
13	499.67	102.00	(i)
64	379.67	27.0	(i)
65	379.67	57.0	(i)

Table 5.16. Vapor Composition Data Records that Have Been Flagged in Data Set 545

Data Point #	Temp. (R)	Y_1	Criteria for Outliers	Y_2	Criteria for Outliers
11	499.67	0.6213	(i)	0.3787	(i)
40	439.67	0.4570	(i)	0.5430	(i)

Table 5.17. Pressure Data Records that Have Been Flagged in Data Set 501

Data Point #	Temp. (R)	Press. (psia)	Criteria for Outliers
7	559.67	200.0	(i) , (iii)

Table 5.18. Vapor Composition Data Records that Have Been Flagged in Data Set 501

Data Point #	Temp. (R)	Y_1	Criteria for Outliers	Y_2	Criteria for Outliers
7	559.67	0.7027	(i)	0.2973	(i) , (iii)

Discussion

The point-by-point analyses presented above for the two binary systems highlights the data screening methodology adopted for evaluating the use of F-tests to screen the VLE data in the GPA Database. A similar analyses was applied to the rest of the binary VLE systems that have 39 data sets and a ternary VLE system that has 2 data sets. Tables 5.19 and 5.20 present a summary of data records that require further examination. These data records were identified as possible outliers.

As an overview, data records that were flagged for displaying "higher-than-expected" deviations in Data Sets 461, 507, 500, 571, 573 and 641 may be a result of model-lack-of-fit due to the high pressure and temperature conditions. Gross systematic errors can also be found in these data sets.

Most of the results in this methodology indicate that the F-tests in the model-prediction methodology are reliable in screening the VLE data in the GPA Database.

To further analyze and examine the results presented in the data screening methodology, graphical deviation plots should be used. According to Twomey (29), major shortcomings of an EOS can clearly show up on a three-dimensional graph. Also, systematic trends between the model predictions and the experimental data can be analyzed.

Another recommended method is a runs test (26) which is one of the non-parametric tests. By using this statistical method, trends and randomness of the deviations from model predictions can be detected.

Table 5.19. Summary of Possible Outliers in the Binary VLE Systems

DSN	Comp#1	Comp#2	Raw/Smooth	Total Pts	# $P_{flagged}$	# $y(_1)_{flagged}$	# $y(_2)_{flagged}$
447	Methane	n-Hexane	S	57	3	4	4
461	Methane	n-Hexane	Unknown	104	0	7	7
570	Methane	Hydrogen	R	25	2	2	2
385	Methane	Hydrogen	Unknown	15	0	1	1
450	Ethane	Carbon Dioxide	R	15	0	0	0
566	Ethane	Carbon Dioxide	R	17	0	2	2
519	Ethane	Carbon Dioxide	R	15	0	0	0
423	Ethane	Hydrogen Sulfide	Unknown	27	0	1	1
796	Ethane	Hydrogen Sulfide	R	45	2	2	2
341	Propane	Carbon Dioxide	S	15	0	0	0
451	Propane	Carbon Dioxide	R	24	0	2	2
512	Propane	Carbon Dioxide	R	21	0	0	0
440	Propane	n-Pentane	Unknown	38	1	3	3
540	Propane	n-Pentane	R	80	0	6	6
346	Isobutane	Carbon Dioxide	R	33	0	1	1
615	Isobutane	Carbon Dioxide	R	69	2	2	2
350	Isopentane	Cyclohexane	Unknown	7	0	0	0
790	Isopentane	Cyclohexane	R	30	0	0	0
351	Isopentane	Methylcyclohexane	Unknown	7	0	0	0

Table 5.19. Summary of Possible Outliers in the Binary VLE Systems (continued)

DSN	Comp#1	Comp#2	Raw/Smooth	Total Pts	# $P_{flagged}$	# $y(_1)_{flagged}$	# $y(_2)_{flagged}$
791	Isopentane	Methylcyclohexane	R	49	0	1	1
792	n-Heptane	Cyclohexane	R	27	0	1	1
828	n-Heptane	Cyclohexane	R	19	0	1	1
433	n-Heptane	Toluene	R	25	0	1	1
823	n-Heptane	Toluene	R	24	0	0	0
507	Carbon Monoxide	Carbon Dioxide	R	18	2	1	1
557	Carbon Monoxide	Carbon Dioxide	R	38	0	0	0
500	Carbon Monoxide	Hydrogen	R	132	15	9	9
571	Carbon Monoxide	Hydrogen	R	16	1	1	1
349	Carbon Dioxide	Hydrogen Sulfide	R	83	1	3	3
432	Carbon Dioxide	Hydrogen Sulfide	S	83	1	4	4
522	Carbon Dioxide	Nitrogen	R	34	0	0	0
569	Carbon Dioxide	Nitrogen	Unknown	16	0	0	0
546	Hydrogen	Nitrogen	R	13	0	0	0
573	Hydrogen	Nitrogen	Unknown	32	2	0	0

Table 5.20. Summary of Possible Outliers in the Ternary VLE System

DSN	Comp#1	Comp#2	Comp#3	Raw/ Smooth	Total Pts	# P_{flagged}	# $y(_{1})_{\mathrm{flagged}}$	# $y(_{2})_{\mathrm{flagged}}$	# $y(_{3})_{\mathrm{flagged}}$
641	Hydrogen	Methane	Carbon Monoxide	R	196	13	4	11	1
644	Hydrogen	Methane	Carbon Monoxide	Unknown	38	3	0	1	0

CHAPTER VI

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary

In this work, same-components data sets reported by different investigators in the VLE systems of the GPA Database were evaluated. These VLE data sets mainly include the measurements of pressure, P; temperature, T; component mole fraction in the liquid phase, x; and component mole fraction in the vapor phase, y. Error analyses based on the SRK EOS and F-tests were employed. Methodologies included in these error analyses are population-check methodology, model-predictions-check methodology, and data screening methodology.

In the population-check methodology, F-tests were used to determine whether there are major differences in the same-components data sets reported by different investigators. In the model-predictions-check methodology, F-tests were used to determine whether

(i) The model predictions of P and y are reasonable.

(ii) The SRK EOS predicted equally well in the data sets that came from the same populations.

The calculated values of P and y from the SRK EOS were compared to the corresponding experimental data in the data screening methodology. By checking the percentage differences of P and deviations of y point-by-point, reliability of F-tests in the model-prediction-check methodology was determined. Also, data discrepancies and data meriting further examination were identified. About 1,800 data points in 44 VLE data sets were tested under these error analyses.

Conclusions

For this study, only a limited number of VLE systems in the GPA Database can be tested by the thermodynamic consistency tests based on Gibbs-Duhem equation, since the experimental volumetric and enthalpy data are not available at the required temperature and pressure conditions.

Based on the results of population-check methodology, variables in the data sets of more than 65% of 19 VLE systems are expected to come from the same populations. Therefore, major differences are not expected in the measurements of those variables in the data sets. For the variables that did not come from the same populations, different investigated temperature and pressure ranges are considered to be the main factors for the major differences.

The results in the model-predictions-check methodology indicate that model predictions are reasonable in almost all the selected VLE data sets. Also, the SRK EOS did predict equally well in most of the data sets that came from the same populations. Therefore, the results of the model-predictions-check methodology support F-tests in the population-check methodology.

Most of the results in the data screening methodology indicate that F-tests in the model-prediction-check methodology are reliable. Thus, F-tests are considered an effective tool to screen the VLE data in the GPA Database. About 3 % pressure data records and 4 % vapor composition data records were identified to require further examination.

Recommendations

For future work, graphical deviation plots (29) or run tests (26) should be used to further analyze and examine the results presented in the data screening methodology. By using either one of these methods, systematic trends between the model predictions and the experimental VLE data can be analyzed.

LITERATURE CITED

1. Bertucco, A., Barolo, M. and Elvassore, N., "Thermodynamic Consistency of Vapor-Liquid Equilibrium Data at High Pressure," *AICHE J.*, 43, 547-554 (1997).

2. Cho, T. H., Ochi, K. and Kojima, K., "Measurement of Liquid Vapor Equilibrium for Systems with Limited Miscibility," *Fluid Phase Equilibria*, 11, 137-143 (1983).

3. Christiansen, L. J. and Fredenslund, A., "Thermodynamic Consistency Using Orthogonal Collocation or Computation of Equilibrium Vapor Compositions at High Pressure," *AICHE J.*, 21, 49-57 (1975).

4. Dohnal, V. and Fenclova, D., "A new Procedure for Consistency Testing of Binary Vapor Liquid Equilibrium Data," *Fluid Phase Equilibia*, 21, 211-215 (1985).

5. Erbar, J. H., "The GPA*Sim Program," Gas Processors Association, Tulsa, Oklahoma (1980).

6. Fredenslund, A., Gmehling, J. and Rasmussen, P., "*Vapor-Liquid Equilibria Using UNIFAC*," Elsevier, Amsterdam, 68 (1977).

7. Gasem, K. A. M., Class-notes, "Principle of Thermodynamics (Cheng 5843)", Chemical Engineering Department of Oklahoma State University, (1998).

8. Hala, E., "The Liquid Equilibrium. XXI. The Thermodynamic Consistency of System with Limited Miscibility in the Liquid Phase," *Czech. Chem. Commun.*, 25, 394 (1960).

9. Herington, E. f. G, "Test for The Consistency of Experimental Isobaric Vapor-Liquid Equilibrium Data," *J.Inst.Petrol.*, 37, 457 (1951).

10. Kojima, K., Moon, H. M., and Ochi, K., "Thermodynamic Consistency of Vapor-Liquid Equilibrium Data--Methanol~Water, Benzene~Cyclohexane and Ethyl Methyl Ketone~Water," *Fluid Phase Equilibria*, 56, 269-284 (1990).

11. Kollar-Hunek, K., Kemeny, S., Heberger, K., Angyal, P. and Thury, E., "Thermodynamic Consistency Test for Binary VLE Data," *Fluid Phase Equilibria*, 27, 405-412 (1986).

12. Liebermann, E. and Fried, V., "Thermodynamic Consistency Test Methods," *Ind. Eng. Chem. Fundam.*, 11, 280-288 (1972).

13. Maase, Eric, Personnel Communication, Oklahoma State University, 11/20/98.

14. McDermott, C. and Ellis, S. R. M., "A Multicomponent Test," *Chemical Engineering Science*, 20, 293-296 (1965).

15. Mixon, F. O., B. Gumowski, and B. H. Carpenter, "Computation of Vapor-Liquid Equilibrium Data from Solution Vapor Pressure Measurements," *Ind. Eng. Chem. Fundamentals*, 4, 455 (1965).

16. Moon, H. M., Ochi, K. and Kojima, K., "Thermodynamic Consistency of Vapor-Liquid Equilibrium Data--Alcohol-Hydrocarbon Systems," *Fluid Phase Equilibria*, 62, 29-40 (1991).

17. Moser, Barry, Personnel Communication, Oklahoma State University, 2/8/99.

18. Ozokwelu, E. D., "Development of A Modified Soave-Redlich-Kwong Equation of State," M. S. Thesis, Oklahoma State University, (1981).

19. Prausnitz, J. M., Lichtenthaler, R. N. and Azevedo, E. G. d., "*Molecular Thermodynamics of Fluid-Phase Equilibria*," 2nd Edition, P T R Prentice Hall, New Jersey, 221-225 (1986).

20. Raal, J. D., Code, R. K. and Best, D. A., "Examination of Ethanol-N-Heptane, Methanol-N-Hexane Systems Using New Vapor-Liquid Equilibrium Still.," *J. Chem. Eng. Data.*, 17, 211 (1972).

21. Rastogi, A., "Evaluation And Maintenance of An Enthalpy Database," M. S. Thesis, Oklahoma State University, (1996).

22. Redlich, O. and Kister, A. T., "Algebraic Representation of Thermodynamic Properties and Classification of Solutions," *Ind. Eng. Chem.*, 40, 345 (1948).

23. Samuels, M. R., "Interpreting Thermodynamic Consistency: How Bad Is 'Bad'?," *Ind. Eng. Chem. Fundam.*, 11, 422 (1972).

24. Samuels, M. R., Ulrichson, D. L. and Stevenson, F. D., "Interpretation of Overall Area Tests for Thermodynamic Consistency: The Effect of Random Area," *AICHE J.*, 18, 1004-1009 (1972).

25. Sanghavi, P., "Design and Development of A Thermodynamic Properties Database Using The Relational Data Model," M. S. Thesis, Oklahoma State University, (1995).

26. Sen, P. K. and Krishnaiah, P. R., "*Non-parametric Methods*," North-Holland Amsterdam, New York, Oxford, 4, 107 (1984).

27. Soave, G., "Equilibrium Constants from A Modified Redlich-Kwong Equation of State," *Chemical Engineering Science*, 27, 1197 (1972).

28. Stevenson, F. D. and Sater, V. E., " Local Thermodynamic Consistency of Vapor-Liquid Equilibrium Data For Binary and Multicomponent Systems," *AICHE J.*, 12, 586 (1966).

29. Twomey, D. W., "Evaluation and a New Reporting Method for Enthalpy Data," M. S. Thesis, Oklahoma State University, (1998).

30. Ulrichson, D. L. and Stevenson, F. D., "Effects of Experimental Errors on Thermodynamic Consistency and on Representation of Vapor-Liquid Equilibrium Data," *Ind. Eng. Chem. Fundam.*, 11, 287 (1972).

31. Van Ness, H. C., "Thermodynamics in the treatment of (Vapor + Liquid) equilibria," *J. Chem. Thermodynamics*, 27, 113 (1995).

32. Van Ness, H. C., Byer, S. M. and Gibbs, R.E., "An Appraisal of Data Reduction Method," *AICHE J.*, 19, 238-244 (1973).

33. Van Ness, H. C., Byer, S. M. and Gibbs, R.E., "Precise Testing of Binary Vapor-Liquid Equilibrium Data by The Gibbs-Duhem Equation," *Chemical Engineering Science*, 11, 118 (1959).

34. Volk, William, "*Applied Statistics for Engineers*," McGraw-Hill Book Company, Inc., (1958).

35. Wankat, P. C., "*Equilibrium Staged Separations*," Elsevier, New York, (1988).

36. Wellendorf, G. R., "K-Values of Polar Mixtures Predicted By The Soave-Redlich-Kwong Equation of State," M.S. Thesis, Oklahoma State University, (1972).

37. Wisniak, J., "The Herington Test for Thermodynamic Consistency," *Ind. Eng. Chem. Res.*, 33, 177-180 (1994).

38. Wisniak, J., "The New Test for the Thermodynamic Consistency of vapor-Liquid Equilibrium," *Ind. Eng. Chem. Res.*, 32, 1531-1533 (1993).

APPENDIX A

PROCEDURES OF CREATING MATRIX SYSTEMS FOR VAPOR-LIQUID-EQUILIBRIUM (VLE) AND DEW-POINT AND BUBBLE-POINT DATA IN THE GPA DATABASE

The version of the GPA Database created in Access version 2.0 was used throughout the process of creating 6 matrix systems in this document. As mentioned earlier, theses matrix systems were used to identify the components in binary, ternary, and multi-component systems for VLE and dew-point and bubble-point data. Access 2.0 provides many advantages to create, maintain, and manipulate records and files in the GPA Database.

Since relational tables were used to store data records in the GPA Database, several steps were taken to retrieve the desired records and files. They are:

(i) Identify data records that are stored in the tables of the GPA Database.

(ii) Extract data information using select queries and cross-tab queries, and export data to Microsoft Excel.

(iii) Create six maps by using Microsoft Excel based on the results of the created queries.

(iv) Check the information of the created queries and maps to avoid human and computer errors.

Step 1. Identify Data Records that are Stored in Tables of the GPA Database

To retrieve the desired information efficiently, one has to know what values or data records are stored in the tables. Many abbreviations were used for the headings of the table columns and viewing the tables themselves is necessary. Maase (34) helped build this database, and provided information in this step.

Step 2. Data Extraction

Select queries, which are defined as new queries, were widely used. The design window of the select query allows one to add the required tables or queries and select the desired fields. The following table shows the names of the added tables and the selected fields used to begin with the design of the first six queries:

Table A.1. Names of Added Tables and the Selected Fields

Added Tables' Name	Selected Fields
DS_COMP	DSN (Data Set Number), CID (Component Identity Number)
WProp	CNAME (Component Name)
DS	DT (Data Type), NC (Number of Component), NDP (Number of Data Points), T_{min} (Minimum Temperature), T_{max} (Maximum Temperature), P_{min} (Minimum Pressure), P_{max} (Maximum Pressure)

To improve the search of these queries, the sorting field of DSN was set to ascending order, and certain criteria are required to set DT and NC. Since 6 different matrix systems have been created, the required criteria can be found in Table A.2.

Table A.2. Criteria Used in the Queries for Creating Different Matrix Systems

Matrix Systems	DT (data type)	NC (number of component)
Binary VLE	= 2	= 2
Binary Dew-Point and Bubble-Point	= 1	= 2
Ternary VLE	= 2	= 3
Ternary Dew-Point and Bubble-Point	= 1	= 3
Multi-VLE	= 2	> 3
Multi- Dew-Point and Bubble-Point	= 1	> 3

The required information for these systems listed above could be obtained when the "run" button on the menu bar was clicked.

Among these matrix systems, the matrix systems for binary VLE, dew-point and bubble-point systems were considered more complexes to design and build. The size of the binary VLE systems is largest compared to the other systems. Thus, a three-query method was required to extract the data in binary systems.

In this method, two created queries mentioned above are considered as the first queries to extract data in binary VLE, and dew-point and bubble-point systems. Another two select queries were created to identify the components that are pairing up in one data set. The second queries had the same settings as in the first query, except for the criteria of CNAME or CID. The criteria of these two entries need to set to one particular component in the binary systems, for example n-Butane. Then, create the third query based on the first two queries and enter the following settings:

Table A.3. Names of Added Queries, Selected Fields and Criterion

Added Queries	Selected Fields	Criterion
First query that listed all the components in a binary system	CNAME, CID, NDP, T_{min}, T_{max}, P_{min}, P_{max}	Set the DSN of the first query equal to the second Query
Second query that listed only one component in a binary system	CNAME	

In the data-sheet view of a third query shown in Table A.4., a group of data sets that have one identical component was obtained. Thus, two components were clearly identified in one data set. This three-query-method was then applied to the rest of the components in the binary systems. Eventually, a group of third queries were created and all the data sets in binary systems with the desired information were collected.

All the third queries were then exported to Microsoft Excel and grouped into a single file by using copy and paste functions. Therefore, binary VLE and binary dew-point and bubble-point systems would each have a single Excel file that could be imported back to Access 2.0. With the new imported tables, the data sets that have same components were further identified using the cross-tab queries.

Cross-tab query is a special query that can "map" a table in a spread-sheet-like format. This type of query was also applied to the ternary and multi-component systems after the data in ternary and multi-component systems were extracted by the first query.

Table A.4. Results of the Third Query Designed for Binary VLE Systems

DSN	Comp 1	Comp 2	CID	T(min)	T(max)	P(min)	P(max)	#Pt
340	n-Butane	Nitrogen	60	559.67	759.67	518.0	4219.0	27
378	n-Butane	n-Heptane	14	627.67	952.67	100.0	400.0	6
386	n-Butane	Hydrogen	58	589.77	709.65	403.0	2447.7	60
398	n-Butane	Carbon Dioxide	53	559.67	739.67	51.5	1184.0	71
413	n-Butane	Nitrogen	60	559.67	739.67	236.0	3402.0	34
419	n-Butane	Carbon Dioxide	53	410.37	509.67	4.8	599.5	29
424	n-Butane	Nitrogen	60	509.77	609.57	82.0	2009.0	13
452	n-Butane	Carbon Dioxide	53	491.67	491.67	15.3	505.5	15
588	n-Butane	Nitrogen	60	559.67	679.67	101.0	3000.0	10
616	n-Butane	Carbon Dioxide	53	556.38	709.67	48.7	1153.2	56
619	n-Butane	Nitrogen	60	610.47	684.63	180.0	3206.0	31
405	n-Butane	n-Decane	26	559.67	919.67	0.1	714.0	78
812	n-Butane	Carbon Dioxide	53	491.67	491.67	16.8	30.4	18

For a cross-tab query to work, one must specify at least one field to be a row heading, one field to be a column heading, and one field to be a value in the query. Thus, row headings, column headings, and values in the cross-tab queries were specified as shown in Table A.5:

Table A.5. Specifications of the Crosstab Queries for Binary, Ternary and Multi-component Systems

Matrix Systems	Row Heading(s)	Column Heading	Values
Binary	CNAME (second component of the binary system), T_{min}, T_{max}, P_{min}, P_{max}, NDP, DSN	CNAME (first component of the binary system)	CID
Ternary	CNAME, CID	DSN	DSN
Multi-component	CNAME, CID	DSN	DSN

Tables A.6, A.7 and A.8 present the results for binary, ternary and multi-component dew-point and bubble-point systems. The empty spaces of the tables signify missing components in the particular data set. Also, data sets that have the same components can easily be identified in these tables, for example, Data Set 758, 759 and 760 are the same-component data sets in Table A.8.

Table A.6. Crosstab Query Result of Binary DB Systems

Comp 2	T(min)	T(max)	P(min)	P(max)	#Pt	DSN	Carbon Dioxide	Hydrogen Sulfide	Methane	Propane
Carbon Dioxide	275.67	394.67	84.4	939.0	110	1			53	
Carbon Dioxide	379.67	474.17	190.0	1212.0	18	2			53	
Carbon Dioxide	275.67	394.67	169.5	938.0	140	8			53	
Hydrogen	558.47	659.57	1000.0	8000.0	14	766				58
n-Butane	259.66	499.67	19.8	1865.0	173	3			5	
n-Heptane	419.67	499.67	19.8	3272.0	53	6			14	
n-Hexane	342.90	491.69	19.9	2675.0	114	5			9	
n-Pentane	311.71	491.69	20.1	2200.0	118	4			7	
n-Pentane	317.18	491.69	250.0	2190.0	23	7			7	
Nitrogen	398.07	477.47	200.0	1940.0	7	9	60			
Nitrogen	399.87	512.77	100.0	1470.0	5	10		60		
Toluene	339.67	499.67	50.0	7070.0	85	11			43	

Table A.7. Crosstab Query Result of Ternary DB Systems

Component	CID	12	13	14	15	16	5000
Carbon Dioxide	53	12	13	14		16	
Ethane	2	12					
Hydrogen Sulfide	54		13		15	16	
Methane	1	12	13	14	15	16	5000
Nitrogen	60			14	15		
Propane	3						5000
Water	62						5000

Table A.8. Crosstab Query Result of Multi-DB Systems

Component	CID	17	18	19	20	21	758	759	760
Carbon Dioxide	53	17			20	21			
Ethane	2		18	19	20	21	758	759	760
Hydrogen Sulfide	54	17							
Methane	1	17	18	19	20	21	758	759	760
n-Butane	5			19	20	21	758	759	760
n-Heptane	14					21			
n-Hexane	9				20	21			
n-Octane	22				20	21			
n-Pentane	7			19	20	21	758	759	760
Nitrogen	60		18	19	20	21	758	759	760
Propane	3		18	19	20	21	758	759	760
Water	62	17							

Step 3. Maps Drawing

After collecting all the necessary information from the GPA Database, two different methods were used to draw six maps in Excel. Refer to the figures of Chapter III in this document for more details.

Step 4. Data Confirmation

Since data extraction required a lot of steps, selected checks for the information in the created queries and maps were necessary. This process can detect the computer errors when using Access 2.0 and Microsoft Excel, and human errors throughout the first three steps of these procedures.

To carry out the selected checks, new queries were needed to check some of the retrieved data from the queries that have been created in Step 2. The settings of these new queries are the same as shown in Tables A.1 and A.2. However, criterion of the data set number (DSN) was set to a data set that was desired to check.

A few data sets were checked for a created query to confirm whether the extracted information is correct. For the maps that were created in Excel, one can also use the same approach to check the available components in certain data sets.

APPENDIX B

SAMPLE OF AN INPUT FILE

FOR THE FORTRAN PROGRAM

```
Reference Index:              569
Title, Source, Page and Date Published
Vapor-Liquid Equilibria in the Nitrogen-Methane-n-Butane
System
Tex. Univ., Austin, Ph.D. Thesis, 57 pp, 1959
 1  Total Authors
      Sauer, R.N.
Data Set Index:  587          Ref. Index: 569
Data Type - 2 - Vapor Liquid Equilibrium Data
#Points  #Comps  MaxT(F)  MinT(F)  MaxP(psia)  MinP(psia)
 4        2        559.67    559.67    111         1803
 1     Methane
 2     n-Butane
****** VAPOR-LIQUID EQUILIBRIA DATA SET EXPORT ******
 Raw/Smooth/Unknown- Null  Data Source- G

DataPt Temp(F) Press(psia) DataFit Consistency CriticalPt
 1      559.67     111.0         Null        Null       Null 1
   Comp      X             Y             Z
     1     0.02          0.4945        0
     2     0.98          0.5055        0
DataPt Temp(F) Press(psia) DataFit Consistency CriticalPt
 2      559.67     499.0         Null        Null       Null 2
   Comp      X             Y             Z
     1     0.1517        0.8491        0
     2     0.8483        0.1509        0
DataPt Temp(F) Press(psia) DataFit Consistency CriticalPt
 3      559.67     1017.0        Null        Null       Null 3
   Comp      X             Y             Z
     1     0.3185        0.8918        0
     2     0.6815        0.1082        0
DataPt Temp(F) Press(psia) DataFit Consistency CriticalPt
 4      559.67     1803.0        Null        Null       Null 4
   Comp      X             Y             Z
     1     0.5928        0.8166        0
     2     0.4072        0.1834        0
```

APPENDIX C

TABLES OF F-TESTS FOR THE POPULATION-CHECK METHODOLOGY APPLIED IN VLE SYSTEMS

Table C.1. F-Tests for the Population-Check methodology Applied to Binary VLE Systems

Case #	Data Set #	Comp#1	Comp#2	Tests	T	P	x_1	x_2	y_1	y_2
1	447	Methane	n-Hexane	Fcalculated	2.63	1.59	4.17	4.17	445.97	445.92
	461			$F_{0.05}$	1.43	1.50	1.50	1.50	1.43	1.43
				H_o	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected
				$F_{0.01}$	1.71	1.78	1.78	1.78	1.71	1.71
				H_o	Rejected	Accepted	Rejected	Rejected	Rejected	Rejected
2	385	Methane	Hydrogen	Fcalculated	11.88	3.07	14.38	14.38	127.13	128.49
	570			$F_{0.05}$	2.35	2.35	2.35	2.35	2.35	2.35
				H_o	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected
				$F_{0.01}$	3.43	3.43	3.43	3.43	3.43	3.43
				H_o	Rejected	Accepted	Rejected	Rejected	Rejected	Rejected
3a	450	Ethane	Carbon Dioxide	Fcalculated	0.11	3.59	1.20	1.20	1.30	1.30
	566			$F_{0.05}$	2.44	2.44	2.44	2.44	2.44	2.44
				H_o	Accepted	Rejected	Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	3.62	3.62	3.62	3.62	3.62	3.62
				H_o	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
3b	519	Ethane	Carbon Dioxide	Fcalculated	NV	4.13	1.02	1.02	1.12	1.12
	566			$F_{0.05}$		2.44	2.44	2.44	2.44	2.44
				H_o		Rejected	Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$		3.62	3.62	3.62	3.62	3.62
				H_o		Rejected	Accepted	Accepted	Accepted	Accepted

Note: NV means F-ratio cannot be found in this variable because one of the data sets is isothermal and so the variance is equal to zero.

Table C.1. F-Tests for the Population-Check methodology Applied to Binary VLE Systems (continued)

Case #	Data Set #	Comp#1	Comp#2	Tests	T	P	x_1	x_2	y_1	y_2
4	423	Ethane	Hydrogen Sulfide	Fcalculated	1.08	1.12	1.66	1.66	1.44	1.44
	796			$F_{0.05}$	1.83	1.83	1.83	1.83	1.83	1.83
				H_o	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	2.38	2.38	2.38	2.38	2.38	2.38
				H_o	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
5a	341	Propane	Carbon Dioxide	Fcalculated	2.61	1.10	1.18	1.18	1.48	1.48
	451			$F_{0.05}$	2.14	2.36	2.14	2.14	2.14	2.14
				H_o	Rejected	Accepted	Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	2.97	3.44	2.97	2.97	2.97	2.97
				H_o	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
5b	341	Propane	Carbon Dioxide	Fcalculated	2.15	2.34	2.48	2.48	11.55	11.55
	512			$F_{0.05}$	2.23	2.23	2.23	2.23	2.23	2.23
				H_o	Accepted	Rejected	Rejected	Rejected	Rejected	Rejected
				$F_{0.01}$	3.13	3.13	3.13	3.13	3.13	3.13
				H_o	Accepted	Accepted	Accepted	Accepted	Rejected	Rejected
5c	451	Propane	Carbon Dioxide	Fcalculated	1.22	2.58	2.10	2.10	7.82	7.82
	512			$F_{0.05}$	2.04	2.09	2.09	2.09	2.09	2.09
				H_o	Accepted	Rejected	Rejected	Rejected	Rejected	Rejected
				$F_{0.01}$	2.78	2.88	2.88	2.88	2.88	2.88
				H_o	Accepted	Accepted	Accepted	Accepted	Rejected	Rejected

Table C.1. F-Tests for the Population-Check methodology Applied to Binary VLE Systems (continued)

Case #	Data Set #	Comp#1	Comp#2	Tests	T	P	x_1	x_2	y_1	y_2
6	440	Propane	n-Pentane	Fcalculated	18.84	2.34	1.13	1.13	2.57	2.57
	540			$F_{0.05}$	1.56	1.56	1.56	1.56	1.56	1.56
				H_o	Rejected	Rejected	Accepted	Accepted	Rejected	Rejected
				$F_{0.01}$	1.85	1.85	1.85	1.85	1.85	1.85
				H_o	Rejected	Rejected	Accepted	Accepted	Rejected	Rejected
7	346	Isobutane	Carbon Dioxide	Fcalculated	1.27	1.13	1.12	1.12	1.12	1.12
	615			$F_{0.05}$	1.62	1.70	1.62	1.62	1.62	1.62
				H_o	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	1.99	1.99	1.99	1.99	1.99	1.99
				H_o	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
8	350	Isopentane	Cyclohexane	Fcalculated	1.09	NV	1.26	1.26	1.03	1.03
	790			$F_{0.05}$	3.82		3.82	3.82	3.82	3.82
				H_o	Accepted		Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	7.24		7.24	7.24	7.24	7.24
				H_o	Accepted		Accepted	Accepted	Accepted	Accepted
9	351	Isopentane	Methylcyclohexane	Fcalculated	5.44	NV	1.62	1.62	2.07	2.07
	791			$F_{0.05}$	3.75		2.30	2.30	2.30	2.30
				H_o	Rejected		Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	7.10		3.20	3.20	3.20	3.20
				H_o	Accepted		Accepted	Accepted	Accepted	Accepted

Note: NV means F-ratio cannot be found in this variable because one of the data sets is isobaric and so the variance is equal to zero.

Table C.1. F-Tests for the Population-Check methodology Applied to Binary VLE Systems (continued)

Case #	Data Set #	Comp#1	Comp#2	Tests	T	P	x_1	x_2	y_1	y_2
10	792	n-Heptane	Cyclohexane	Fcalculated	1.07	NV	1.08	1.08	1.09	1.09
	828			$F_{0.05}$	2.14		2.14	2.14	2.14	2.14
				H_o	Accepted		Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	2.97		2.97	2.97	2.97	2.97
				H_o	Accepted		Accepted	Accepted	Accepted	Accepted
11	433	n-Heptane	Toluene	Fcalculated	1.09	NV	1.14	1.14	1.16	1.16
	823			$F_{0.05}$	2.04		2.04	2.04	2.04	2.04
				H_o	Accepted		Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	2.78		2.78	2.78	2.78	2.78
				H_o	Accepted		Accepted	Accepted	Accepted	Accepted
12	507	Carbon Monoxide	Carbon Dioxide	Fcalculated	1.14	1.72	1.73	1.73	1.61	1.61
	557			$F_{0.05}$	1.85	1.98	2.08	2.08	2.00	2.00
				H_o	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	2.41	2.65	2.85	2.85	2.71	2.71
				H_o	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
13	500	Carbon Monoxide	Hydrogen	Fcalculated	1.00	10.67	2.40	2.39	1.38	1.38
	571			$F_{0.05}$	1.61	1.81	1.12	1.12	1.63	1.63
				H_o	Accepted	Rejected	Rejected	Rejected	Accepted	Accepted
				$F_{0.01}$	1.99	2.35	2.94	2.94	1.98	1.98
				H_o	Accepted	Rejected	Accepted	Accepted	Accepted	Accepted

Note: NV means F-ratio cannot be found in this variable because one of the data sets is isobaric and so the variance is equal to zero.

Table C.1. F-Tests for the Population-Check methodology Applied to Binary VLE Systems (continued)

Case #	Data Set #	Comp#1	Comp#2	Tests	T	P	x_1	x_2	y_1	y_2
14	349	Carbon Dioxide	Hydrogen Sulfide	Fcalculated	1.76	1.66	1.43	1.43	1.39	1.39
	432			$F_{0.05}$	1.44	1.44	1.44	1.44	1.44	1.44
				H_o	Rejected	Rejected	Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$	1.68	1.68	1.68	1.68	1.68	1.68
				H_o	Rejected	Accepted	Accepted	Accepted	Accepted	Accepted
15	522	Carbon Dioxide	Nitrogen	Fcalculated	NV	1.30	1.04	1.04	1.04	1.04
	569			$F_{0.05}$		2.23	2.23	2.23	1.97	1.97
				H_o		Accepted	Accepted	Accepted	Accepted	Accepted
				$F_{0.01}$		3.16	3.16	3.16	2.60	2.60
				H_o		Accepted	Accepted	Accepted	Accepted	Accepted
16	546	Hydrogen	Nitrogen	Fcalculated	21.49	15.24	15.22	15.22	1.43	1.43
	573			$F_{0.05}$	2.20	2.20	2.20	2.20	2.20	2.20
				H_o	Rejected	Rejected	Rejected	Rejected	Accepted	Accepted
				$F_{0.01}$	3.10	3.10	3.10	3.10	3.10	3.10
				H_o	Rejected	Rejected	Rejected	Rejected	Accepted	Accepted

Note: NV means F-ratio cannot be found in this variable because one of the data sets is isothermal and so the variance is equal to zero.

Table C.2. F-Tests for the Population-Check methodology Applied to a Ternary VLE System

Case #	Data Set #	Comp#1	Comp#2	Comp#3	Tests	T	P	x_1	x_2	x_3	y_1	y_2	y_3
1	641	Hydrogen	Methane	Carbon	Fcalculated	5.61	5.94	5.15	1.17	1.14	1.55	7.30	1.21
	644			Monoxide	$F_{0.05}$	1.58	1.58	1.58	1.58	1.58	1.58	1.58	1.47
					H_o	Rejected	Rejected	Rejected	Accepted	Accepted	Accepted	Rejected	Accepted
					$F_{0.01}$	1.92	1.92	1.92	1.92	1.92	1.92	1.92	1.72
					H_o	Rejected	Rejected	Rejected	Accepted	Accepted	Accepted	Rejected	Accepted

Figures
One
and Four.

COMPONENT	Methane	Ethane	Propane	Isobutane	n-Butane	Isopentane	n-Pentane	Neopentane	n-Hexane	n-Heptane	n-Octane	Isooctane	2,2,4-Trimethylpentane	n-Nonane	n-Decane	n-Dodecane	n-Hexadecane	Methylcyclopentane	Cyclohexane	Methylcyclohexane	Ethylcyclohexane	n-Propylcyclohexane	Ethene	Propene	1-Butene	Isobutene	Benzene	Toluene	Ethylbenzene	m-Xylene	p-Xylene	1,3,5-Trimethylbenzene	n-Propylbenzene	1,2,3,4-Tetrahydronapthalene	Phenol	Trietheleneglycol	Carbon Monoxide	Carbon Dioxide	Hydrogen Sulfide	Sulfur Dioxide	Dimethylsulfide	Methylethylsulfide	t-Butlymercaptan	Isopropylmercaptan	Ethylmercaptan	Ammonia	Hydrogen	Oxygen	Nitrogen	Water	Helium	Argon
Methane																																																				
Ethane	X																																																			
Propane	X	X																																																		
Isobutane	X	X	X																																																	
n-Butane	X	X	X	X																																																
Isopentane	X																																																			
n-Pentane	X	X	X																																																	
Neopentane	X																																																			
n-Hexane	X		X																																																	
n-Heptane	X	X	X		X																																															
n-Octane	X																																																			
Isooctane										X																																										
2,2,4-Trimethylpentane																																																				
n-Nonane																																																				
n-Decane	X	X	X		X																																															
n-Dodecane																																																				
n-Hexadecane																																																				
Methylcyclopentane						X			X																																											
Cyclohexane	X	X	X			X			X	X																																										
Methylcyclohexane		X				X			X	X		X																																								
Ethylcyclohexane	X																																																			
n-Propylcyclohexane		X	X																																																	
Ethene	X																																																			
Propene		X		X																																																
1-Butene				X																																																
Isobutene				X																																																
Benzene										X	X	X							X	X																																
Toluene	X	X							X	X		X							X								X																									
Ethylbenzene											X																																									
m-Xylene	X	X																									X	X																								
p-Xylene											X																																									
1,3,5-Trimethylbenzene	X	X	X																																																	
n-Propylbenzene		X	X																																																	
1,2,3,4-Tetrahydronapthalene																																																				
Phenol																																																				
Triethelenеglycol																																																				
Carbon Monoxide	X	X	X								X				X												X																									
Carbon Dioxide	X	X	X	X	X	X	X			X					X		X		X	X	X		X	X	X		X	X		X		X		X	X	X	X															
Hydrogen Sulfide	X	X	X	X		X				X				X						X	X	X						X		X		X					X	X														
Sulfur Dioxide	X																																																			
Dimethylsulfide			X																																																	
Methylethylsulfide			X																																																	
Ethylmercaptan			X																																																	
Isopropylmercaptan			X																																																	
t-Butlymercaptan			X																																																	
Ammonia																																			X																	
Hydrogen	X	X	X	X	X	X					X		X			X			X				X				X								X		X	X														
Oxygen																																						X														
Nitrogen	X	X	X	X	X	X	X								X				X	X	X		X	X				X		X		X					X	X	X	X							X					
Water	X					X																															X	X	X							X	X		X			
Helium	X	X	X																																		X	X									X		X			
Argon	X																																				X	X														

Note: The Xs identify the binary VLE systems

Figure 1. Binary Vapor-Liquid-Equilibrium Systems in the GPA Database

	COMPONENT																											
Data Set Number (DSN)	Methane	Ethane	Propane	n-Butane	Isopentane	n-Pentane	Neopentane	n-Hexane	n-Heptane	3-Ethylpentane	n-Nonane	n-Decane	Ethene	Propene	Toluene	1,3,5-Trimethylbenzene	Phenol	Carbon Monoxide	Carbon Dioxide	Hydrogen Sulfide	Dimethylsulfide	Ethylmercaptan	Ammonia	Hydrogen	Oxygen	Nitrogen	Water	Helium
425	X																		X	X								
621	X		X																					X				
622	X	X				X																						
623	X		X			X																						
624		X	X																							X		
625			X											X												X		
626	X	X																						X				
627	X		X			X																						
628	X	X	X																									
629		X		X		X																						
630	X		X	X																								
631	X																		X	X								
632	X																		X	X								
633			X																							X		X
634	X	X								X																		
635	X		X						X																			
636	X	X																								X		
637	X		X						X																			
638	X		X									X																
639	X		X												X													
640																		X						X			X	
641	X																	X						X				
642	X																		X	X								
643	X					X																					X	
644	X																	X						X				
645	X																							X		X		
646	X																	X								X		
647	X																		X							X		
648																		X	X					X				
649																			X						X	X		
650		X	X																X									
651	X	X																	X									
652	X																		X	X								
653	X																		X							X		
654			X											X				X										
655	X	X	X																									
656		X		X				X																				
657	X																	X	X									
658																		X						X		X		
659	X			X																						X		
660			X															X						X				
661																	X		X								X	
662																	X						X				X	
663																	X							X			X	
664			X																		X	X						
763	X													X										X				
765	X	X				X																						
768	X		X			X																						
773	X	X																						X				
774	X	X																								X		
775	X																							X		X		
777	X		X			X																						
781											X									X				X				
782																X				X				X				
793	X					X	X																					
794	X				X	X																						
807	X												X											X				
808	X	X											X															
809	X		X									X																
811	X	X							X																			

Note: The Xs in a row identify components in a data set of ternary VLE system.

Figure 4. Ternary Vapor-Liquid-Equilibrium Systems in the GPA Database

Activities: Phi Lambda Upsilon Honor Society
Dancer at [illegible] Night and Malaysian Night
American Institute of Chemical Engineers
American Chemical Society
Healthcare Information and Management Systems Society

VITA

Sze-Hoong Wong

Candidate for the degree of

Master of Science

Thesis: EVALUATION OF A VAPOR-LIQUID-EQUILIBRIUM DATABASE USING ERROR ANALYSES BASED ON AN EQUATION OF STATE AND STATISTICAL METHODS

Major field: Chemical Engineering

Biographical:

Personal Data: Born in Seremban, Malaysia, on January 23, 1975, the son of Sow-Ha Wong and Sow-Chin Ho.

Education: Received Bachelor of Science degree in Chemical Engineering from Oklahoma State University, Oklahoma, August 1997; completed the requirement for a Master of Science degree at Oklahoma State University in December, 1999.

Experiences: Graduate Research Assistant, Department of Chemical Engineering, Oklahoma State University, 1997-1999. Teaching Assistant, Department of Chemical Engineering, Oklahoma State University, 1997-1998.

Honors: Engineer Intern Certification. (Spring 1997, Oklahoma)
President's List. Received honorable mention of 4.0 GPA (Summer 1997)
Tau Beta Pi National Engineering Honor Society upon graduation
Dean's List. Accomplishment of a GPA 3.62 (Fall, 1996)

Activities: Phi Lambda Upsilon Honor Society
Dancer at Cultural Night and Malaysian Night
American Institute of Chemical Engineers
American Chemical Society
Healthcare Information and Management Systems Society