

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE LORENA

FELIPE RIBEIRO DE SIQUEIRA

**Modelagem termodinâmica do equilíbrio líquido-líquido em sistemas compostos por ácido acético,
água e éster aplicando métodos numéricos e aprendizado de máquina**

Lorena
2022

FELIPE RIBEIRO DE SIQUEIRA

Modelagem termodinâmica do equilíbrio líquido-líquido em sistemas compostos por ácido acético, água e éster aplicando métodos numéricos e aprendizado de máquina

Dissertação apresentada à Escola de Engenharia de Lorena da Universidade de São Paulo para obtenção do título de Mestre em Ciências do Programa de Pós-Graduação em Engenharia Química na área de concentração de Desenvolvimento de Produtos e Processos.

Orientador: Prof. Dr. Leandro Gonçalves de Aguiar

Versão Corrigida

Lorena
2022

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE

Ficha catalográfica elaborada pelo Sistema Automatizado
da Escola de Engenharia de Lorena,
com os dados fornecidos pelo(a) autor(a)

Siqueira, Felipe Ribeiro de
Modelagem termodinâmica do equilíbrio líquido-líquido
em sistemas compostos por ácido acético, água e
éster aplicando métodos numéricos e aprendizado de
máquina / Felipe Ribeiro de Siqueira; orientador
Leandro Gonçalves de Aguiar - Versão Corrigida. -
Lorena, 2022.
62 p.

Dissertação (Mestrado em Ciências - Programa de Pós
Graduação em Engenharia Química na Área de
Desenvolvimento de Produtos e Processos) - Escola de
Engenharia de Lorena da Universidade de São Paulo.
2022

1. Equilíbrio líquido-líquido. 2. Modelo nrtl. 3.
Modelo de florestas aleatórias. 4. Regressão de
parâmetros. 5. Modelagem termodinâmica. I. Título. II.
Aguiar, Leandro Gonçalves de, orient.

RESUMO

SIQUEIRA, F. R. **Modelagem termodinâmica do equilíbrio líquido-líquido em sistemas compostos por ácido acético, água e éster aplicando métodos numéricos e aprendizado de máquina**. 2022. 62 p. Dissertação (Mestrado em Ciências) - Escola de Engenharia de Lorena, Universidade de São Paulo, Lorena, 2022.

A modelagem termodinâmica tem sido uma importante alternativa às práticas experimentais no teste de hipóteses e predição de propriedades de misturas, pois economiza tempo e capital. Neste trabalho, foram experimentados dois modelos para modelagem do equilíbrio de fases líquido-líquido: um modelo de energia de Gibbs em excesso, o modelo NRTL, e um modelo de aprendizado de máquina, regressão por florestas aleatórias. O objetivo foi comparar a performance de ambos na predição da composição das fases existentes na extração do ácido acético de soluções aquosas, compostas por uma mistura envolvendo um acetato, água e ácido acético. Para tanto, foram coletados dados de sistemas ternários compostos por água, ácido acético e por quatro diferentes ésteres, sendo eles os acetatos de butila, isobutila, amila e isoamila. Para o modelo NRTL, foram determinados conjuntos de parâmetros de interação binária por regressão e foi feita a aplicação destes parâmetros na predição das composições das fases em equilíbrio. Já para o modelo de florestas aleatórias, buscou-se um conjunto de hiperparâmetros capazes de treinar os modelos sem produzir superajuste e prever as composições das fases em equilíbrio. Como resultado, foi observado que o modelo de florestas aleatórias possui algumas vantagens sobre o modelo NRTL, sendo possível fazer a predição de composições sem especificar o éster presente no sistema. Para alguns casos aqui expostos, o modelo NRTL não foi capaz de realizar a regressão, resultando em uma composição trivial e distante da experimental.

Palavras-chave: Equilíbrio líquido-líquido. Modelo NRTL. Modelo de florestas aleatórias. Regressão de parâmetros. Modelagem termodinâmica. Extração de ácido acético de soluções aquosas.

ABSTRACT

SIQUEIRA, F. R. **Thermodynamic modeling of liquid-liquid equilibrium in systems composed of acetic acid, water and ester applying numerical methods and machine learning.** 2022. 62 p. Dissertation (Master of Science) - Escola de Engenharia de Lorena, Universidade de São Paulo, Lorena, 2022.

A thermodynamic modeling has an importance in experimental practices in testing alternatives and predicting properties of mixtures, as it saves time and capital. In this one, two working models for liquid-liquid equilibria were tried: NRTL models, an excess Gibbs energy model and a machine learning model called random forest. The objective was compare the performance of both in the prediction of the composition of the phases in the extraction of acetic acid from aqueous solutions using an ester. For that, data from ternary systems were collected for systems composed by four esters: butyl, isobutyl, amyl and isoamyl acetates. For the NRTL model, sets of binary interaction parameters were configured per system and the prediction of the compositions of the phases in equilibrium was made. For the random forest model, a set of hyperparameters was searched, the model was trained and the compositions of the phases was predicted. As a result, it was observed that the random forest model has some advantages over the NRTL model, being possible to make a prediction of compositions without specifying the ester present in the system. For some cases here, the NRTL model was not able to perform the regression and predict the compositions for the phases in equilibrium, resulting in a trivial composition.

Keywords: Liquid-liquid equilibria. NRTL model. Random forest model. Parameter regression. Thermodynamic modeling. Extraction of acetic acid from aqueous solutions.

LISTA DE TABELAS

Tabela 1 – Massas molares e densidades nas condições ambientes dos componentes das misturas estudadas.....	25
Tabela 2 – Características físicas dos dados experimentais relativos à fase aquosa para o sistema água (1) + ácido acético (2) + éster (3) em equilíbrio.....	25
Tabela 3 – Características físicas dos dados experimentais relativos à fase orgânica para o sistema água (1) + ácido acético (2) + éster (3) em equilíbrio.....	26
Tabela 4 – Coeficientes de determinação por temperatura para cada éster estudado.....	28
Tabela 5 – Parâmetros de ajuste obtidos pela regressão do modelo NRTL por conjunto de dados.....	29
Tabela 6 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL.....	29
Tabela 7 – Coeficientes de determinação por fase da regressão e por componente para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias.....	33
Tabela 8 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo NRTL.....	37
Tabela 9 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias.....	39
Tabela 10 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias.....	43
Tabela 11 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetato de butila/isobutila utilizando o modelo de florestas aleatórias.....	48
Tabela 12 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias.....	52

LISTA DE FIGURAS

Figura 1 – Transcrição de uma árvore de decisão utilizada neste trabalho.....	21
Figura 2 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL.....	30
Figura 3 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL.....	30
Figura 4 - Distribuição dos resíduos das frações de acetato de amila na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL.....	30
Figura 5 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL.....	31
Figura 6 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL.....	31
Figura 7 - Distribuição dos resíduos das frações de acetato de amila na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL.....	31
Figura 8 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 298,00K (+ = Experimental, o = Calculado utilizando o modelo NRTL).....	32
Figura 9 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 304,00K (+ = Experimental, o = Calculado utilizando o modelo NRTL).....	32
Figura 10 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 332,00K (+ = Experimental, o = Calculado utilizando o modelo NRTL).....	32
Figura 11 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 366,00K (+ = Experimental, o = Calculado utilizando o modelo NRTL).....	33
Figura 12 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias.....	33
Figura 13 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias.....	34

Figura 14 - Distribuição dos resíduos das frações de acetato de amila na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias.....	34
Figura 15 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias..	35
Figura 16 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias.....	35
Figura 17 - Distribuição dos resíduos das frações de acetato de amila na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias.....	35
Figura 18 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 298,00K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	36
Figura 19 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 304,00K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	36
Figura 20 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 332,00K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	36
Figura 21 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 366,00K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	37
Figura 22 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 293,00 K (+ = Experimental, o = Calculado utilizando o modelo NRTL).....	38
Figura 23 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 298,00 K (+ = Experimental, o = Calculado utilizando o modelo NRTL).....	38
Figura 24 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 313,00 K (+ = Experimental, o = Calculado utilizando o modelo NRTL).....	38

Figura 25 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias.....	39
Figura 26 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias.....	40
Figura 27 - Distribuição dos resíduos das frações de acetato de isobutila na fase aquosa da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias.....	40
Figura 28 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias.....	40
Figura 29 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias.....	41
Figura 30 - Distribuição dos resíduos das frações de acetato de isobutila na fase orgânica da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias.....	41
Figura 31 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 293,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	42
Figura 32 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 298,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	42
Figura 33 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 313,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	42
Figura 34 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias.....	43
Figura 35 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias.....	44

Figura 36 - Distribuição dos resíduos das frações de acetato de amila/isoamila na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias.....	44
Figura 37 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias.....	44
Figura 38 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias.....	45
Figura 39 - Distribuição dos resíduos das frações de acetato de amila/isoamila na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias.....	45
Figura 40 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 293,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	46
Figura 41 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 298,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	46
Figura 42 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 304,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	46
Figura 43 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 313,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	47
Figura 44 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 332,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	47
Figura 45 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 366,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	47
Figura 46 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetato de butila/isobutila utilizando o modelo de florestas aleatórias.....	48

Figura 47 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetato de butila/isobutyla utilizando o modelo de florestas aleatórias.....	49
Figura 48 - Distribuição dos resíduos das frações de acetato de butila/isobutyla na fase aquosa da regressão para o sistema água + ácido acético + acetato de butila/isobutyla utilizando o modelo de florestas aleatórias.....	49
Figura 49 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetato de butila/isobutyla utilizando o modelo de florestas aleatórias.....	49
Figura 50 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetato de butila/isobutyla utilizando o modelo de florestas aleatórias.....	50
Figura 51 - Distribuição dos resíduos das frações de acetato de butila/isobutyla na fase aquosa da regressão para o sistema água + ácido acético + acetato de butila/isobutyla utilizando o modelo de florestas aleatórias.....	50
Figura 52 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de butila/isobutyla na temperatura de 293,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	51
Figura 53 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de butila/isobutyla na temperatura de 298,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	51
Figura 54 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de butila/isobutyla na temperatura de 313,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	51
Figura 55 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias.....	53
Figura 56 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias....	53
Figura 57 - Distribuição dos resíduos das frações de acetatos na fase aquosa da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias.....	53
Figura 58 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias.....	54

Figura 59 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias.....	54
Figura 60 - Distribuição dos resíduos das frações de acetatos na fase orgânica da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias....	54
Figura 61 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 293,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	55
Figura 62 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 298,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	55
Figura 63 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 304,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	56
Figura 64 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 313,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	56
Figura 65 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 332,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	56
Figura 66 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 366,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias).....	57

SUMÁRIO

1 INTRODUÇÃO	15
2 OBJETIVOS	16
3 REVISÃO DA LITERATURA	17
3.1 CONSIDERAÇÕES TERMODINÂMICAS.....	17
3.1.1 Equilíbrio líquido-líquido.....	17
3.1.2 Modelo non-random two liquids (NRTL).....	18
3.2 APRENDIZADO DE MÁQUINA.....	19
3.2.1 Considerações matemáticas e análise.....	20
3.2.2 Florestas aleatórias.....	21
3.3 PROGRAMA UTILIZADO.....	22
4 METODOLOGIA	23
4.1 ESTUDO TERMODINÂMICO DO EQUILÍBRIO DE FASES.....	23
4.1.1 Equacionamento do problema.....	23
4.2 CARACTERÍSTICAS FÍSICAS DOS DADOS E INDICADORES DE QUALIDADE.....	24
4.3 TESTE DE CORRELAÇÃO DE OTHMER-TOBIAS.....	26
5 RESULTADOS E DISCUSSÃO	28
5.1 AVALIAÇÃO DOS DADOS EXPERIMENTAIS.....	28
5.2 RESULTADOS DOS DIFERENTES MODELOS DE REGRESSÃO.....	28
5.2.1 Acetato de amila.....	29
5.2.1.1 Modelo NRTL.....	29
5.2.1.2 Modelo de florestas aleatórias.....	33
5.2.2 Acetato de isobutila.....	37
5.2.2.1 Modelo NRTL.....	37
5.2.2.2 Modelo de florestas aleatórias.....	39
5.2.3 Acetato de amila + Acetato de isoamila.....	43
5.2.3.1 Modelo de florestas aleatórias.....	43
5.2.4 Acetato de butila + Acetato de isobutila.....	48
5.2.4.1 Modelo de florestas aleatórias.....	48
5.2.5 Todos os ésteres.....	52
5.2.5.1 Modelo de florestas aleatórias.....	52
5.3 DISCUSSÃO DOS RESULTADOS.....	57
5.3.1 Modelo NRTL.....	57

5.3.2 Modelo de florestas aleatórias.....	58
6 CONCLUSÕES.....	60
REFERÊNCIAS.....	61

1 INTRODUÇÃO

Segundo Wang *et al.* (2007), o equilíbrio existente entre água, ácido acético e acetatos tem importante aplicação no que se refere à extração de ácido acético de soluções aquosas. Ince e Kirlbaslar (2002) citam que um dos casos onde esta técnica tem sido aplicada é na fermentação e seu estudo é de fundamental importância para que se aumente a recuperação do ácido destas soluções.

Madeira (2019) reforça que água e ácido acético formam uma mistura que se assemelha a um sistema azeotrópico, o que implica em uma grande quantidade de estágios na separação via destilação. Neste caso, o emprego da extração por solvente se faz pertinente, dado seu baixo custo e sua alta eficiência.

Para que se possa compreender o comportamento das fases envolvidas nesta separação de forma rápida e barata, a modelagem termodinâmica tem avançado no que se refere ao número de métodos e à qualidade da predição de propriedades. Este desenvolvimento tem produzido diversas ferramentas, desde métodos numéricos, baseados em modelos matemáticos bem estabelecidos, até métodos de aprendizado de máquina, baseados na busca de padrões nos dados fornecidos ao computador, o que permite à máquina concluir tendências acerca dos dados.

Este trabalho visa estudar o equilíbrio líquido-líquido existente entre água, ácido acético e acetatos com base em dados da literatura, de modo a validar um modelo termodinâmico e comparar seus resultados com modelos mais modernos, baseados em aprendizado de máquina.

2 OBJETIVOS

O objetivo geral deste trabalho é a modelagem do equilíbrio de fases líquido-líquido estabelecido entre ácido acético, água e acetatos por meio do modelo NRTL e de modelo de aprendizado máquina supervisionado, especificamente as florestas aleatórias.

De posse deste objetivo, podemos citar objetivos secundários deste trabalho:

- Determinar um conjunto de parâmetros de interação binária para o modelo NRTL por acetato e avaliar a qualidade da regressão utilizando este modelo;
- Mensurar a qualidade da regressão por meio do modelo de florestas aleatórias;
- Comparar a performance da regressão entre os diferentes modelos e avaliar seus pontos negativos e positivos.

3 REVISÃO DA LITERATURA

3.1 CONSIDERAÇÕES TERMODINÂMICAS

A modelagem termodinâmica é importante pois a obtenção de dados pela via experimental demanda tempo e está atrelada a custos relativos ao experimento. Um modelo termodinâmico capaz de descrever e prever corretamente o comportamento do equilíbrio estudado torna-se vantajoso e pode gerar economia de tempo e capital.

3.1.1 Equilíbrio líquido-líquido

O estado de equilíbrio entre duas fases é determinado por meio de uma igualdade de pressões (equilíbrio mecânico), temperatura (equilíbrio térmico), e fugacidades (equilíbrio químico) de um determinado composto “i” entre as n fases que podem ser formadas (MICHELSEN e MOLLERUP, 2007). Temos, portanto, que a fugacidade f de um composto “i” em um equilíbrio de fases é:

$$f_{i,1} = f_{i,2} = \dots = f_{i,n} \quad (1)$$

Podemos determinar a fugacidade do componente “i” em uma determinada fase em termos do seu coeficiente de atividade:

$$f_i = \gamma_i * P_i^{SAT} * x_i * f_i^0 \quad (2)$$

Sendo γ_i o coeficiente de atividade de “i” na fase em estudo, x_i a fração molar deste componente na fase, P_i^{SAT} a pressão de saturação do composto “i”, calculada pela equação de Antoine, e f_i^0 a fugacidade padrão do estudo.

Logo, em um estado de equilíbrio, temos:

$$\begin{aligned} (\gamma_i * P_i^{SAT} * x_i * f_i^0)_1 &= (\gamma_i * P_i^{SAT} * x_i * f_i^0)_2 = \dots \\ &= (\gamma_i * P_i^{SAT} * x_i * f_i^0)_n \end{aligned} \quad (3)$$

A fugacidade padrão pode ser calculada por:

$$f_i^0 = \phi_i^{SAT} * \exp\left(v_i^{SAT} * \frac{P - P_i^{SAT}}{R * T}\right) \quad (4)$$

Sendo ϕ_i^{SAT} a fugacidade de “i” na saturação, v_i^{SAT} o volume molar de saturação de “i”, P a pressão do sistema, T a temperatura do sistema e R a constante universal dos gases ideais. O termo exponencial é denominado correção de Poynting (POY), o qual considera o efeito de altas pressões sobre as fases em equilíbrio. Sendo assim:

$$\begin{aligned}
& (\mathbf{y}_i * \mathbf{P}_i^{\text{SAT}} * \mathbf{x}_i * \boldsymbol{\phi}_i^{\text{SAT}} * \mathbf{POY}_i)_1 \\
& = (\mathbf{y}_i * \mathbf{P}_i^{\text{SAT}} * \mathbf{x}_i * \boldsymbol{\phi}_i^{\text{SAT}} * \mathbf{POY}_i)_2 = \dots \\
& = (\mathbf{y}_i * \mathbf{P}_i^{\text{SAT}} * \mathbf{x}_i * \boldsymbol{\phi}_i^{\text{SAT}} * \mathbf{POY}_i)_n
\end{aligned} \tag{5}$$

Para pressões baixas, a correção de Poyinting é desprezível, resumindo a equação do equilíbrio a:

$$\begin{aligned}
& (\mathbf{y}_i * \mathbf{P}_i^{\text{SAT}} * \mathbf{x}_i * \boldsymbol{\phi}_i^{\text{SAT}})_1 = (\mathbf{y}_i * \mathbf{P}_i^{\text{SAT}} * \mathbf{x}_i * \boldsymbol{\phi}_i^{\text{SAT}})_2 = \dots \\
& = (\mathbf{y}_i * \mathbf{P}_i^{\text{SAT}} * \mathbf{x}_i * \boldsymbol{\phi}_i^{\text{SAT}})_n
\end{aligned} \tag{6}$$

O coeficiente de atividade é determinado pelos modelos de energia de Gibbs em excesso, tais como Wilson, Margules, Van Laar, NRTL, UNIQUAC ou UNIFAC. Neste trabalho, far-se-á o cálculo do coeficiente pelo modelo NRTL, o qual será descrito no tópico seguinte.

3.1.2 Modelo non-random two liquids (NRTL)

O modelo NRTL (PRAUSNITZ; RENON, 1969) é utilizado para calcular o coeficiente de atividade de um soluto em um solvente líquido para sistemas onde as partes são completamente ou parcialmente solúveis, ajustando os dados experimentais com exatidão significativa.

Nas equações de 7 a 17, x_1 , x_2 e x_3 são as frações molares dos compostos 1, 2 e 3, τ_{ij} e α_{ij} os coeficientes de interação binária do par i-j, tal que i e j variam entre 1 e 3. As equações de 7, 8 e 9 descrevem o modelo NRTL para cálculo dos coeficientes de atividade dos componentes em uma mistura ternária:

$$\begin{aligned}
\ln \gamma_1 &= \frac{\theta_1}{\varphi_1} - \frac{x_1 * \theta_1}{\varphi_1^2} + \frac{x_2 * G_{12}}{\varphi_2} * \left[\tau_{12} - \frac{\theta_2}{\varphi_2} \right] + \frac{x_3 * G_{13}}{\varphi_3} \\
& * \left[\tau_{13} - \frac{\theta_3}{\varphi_3} \right]
\end{aligned} \tag{7}$$

$$\begin{aligned}
\ln \gamma_2 &= \frac{\theta_2}{\varphi_2} - \frac{x_2 * \theta_2}{\varphi_2^2} + \frac{x_1 * G_{21}}{\varphi_1} * \left[\tau_{21} - \frac{\theta_1}{\varphi_1} \right] + \frac{x_3 * G_{23}}{\varphi_3} \\
& * \left[\tau_{23} - \frac{\theta_3}{\varphi_3} \right]
\end{aligned} \tag{8}$$

$$\begin{aligned}
\ln \gamma_3 &= \frac{\theta_3}{\varphi_3} - \frac{x_3 * \theta_3}{\varphi_3^2} + \frac{x_1 * G_{31}}{\varphi_1} * \left[\tau_{31} - \frac{\theta_1}{\varphi_1} \right] + \frac{x_2 * G_{32}}{\varphi_2} \\
& * \left[\tau_{32} - \frac{\theta_2}{\varphi_2} \right]
\end{aligned} \tag{9}$$

Já as equações de 10 a 17 mostram como os coeficientes de interação binária se inserem no cálculo dos parâmetros do modelo NRTL:

$$\boldsymbol{\varphi}_1 = \mathbf{x}_1 + \mathbf{G}_{21} * \mathbf{x}_2 + \mathbf{G}_{31} * \mathbf{x}_3 \quad (10)$$

$$\boldsymbol{\varphi}_2 = \mathbf{x}_2 + \mathbf{G}_{12} * \mathbf{x}_1 + \mathbf{G}_{32} * \mathbf{x}_3 \quad (11)$$

$$\boldsymbol{\varphi}_3 = \mathbf{x}_3 + \mathbf{G}_{13} * \mathbf{x}_1 + \mathbf{G}_{23} * \mathbf{x}_2 \quad (12)$$

$$\boldsymbol{\theta}_1 = \mathbf{x}_2 * \boldsymbol{\tau}_{21} * \mathbf{G}_{21} + \mathbf{x}_3 * \boldsymbol{\tau}_{31} * \mathbf{G}_{31} \quad (13)$$

$$\boldsymbol{\theta}_2 = \mathbf{x}_1 * \boldsymbol{\tau}_{12} * \mathbf{G}_{12} + \mathbf{x}_3 * \boldsymbol{\tau}_{32} * \mathbf{G}_{32} \quad (14)$$

$$\boldsymbol{\theta}_3 = \mathbf{x}_1 * \boldsymbol{\tau}_{13} * \mathbf{G}_{13} + \mathbf{x}_2 * \boldsymbol{\tau}_{23} * \mathbf{G}_{23} \quad (15)$$

$$\mathbf{G}_{ij} = \mathbf{exp}(-\boldsymbol{\alpha}_{ij} * \boldsymbol{\tau}_{ij}) \quad (16)$$

$$\boldsymbol{\tau}_{ij} = \mathbf{a}_{ij} + \frac{\mathbf{b}_{ij}}{\mathbf{T}} \quad (17)$$

O coeficiente α_{ij} é simétrico, ou seja, $\alpha_{ij} = \alpha_{ji}$ e está ligado à não-aleatoriedade presente na mistura, se $\alpha_{ij}=0$ a mistura é completamente aleatória. Para o equilíbrio líquido-líquido e para as faixas de temperatura utilizadas neste trabalho, considerou-se o coeficiente α_{ij} independente da temperatura e fixo em 0,3. Os termos a_{ij} e b_{ij} são parâmetros que ajustam a forma de dependência dos coeficientes τ_{ij} com a temperatura, sendo que, neste caso, estes parâmetros são assimétricos, ou seja, $a_{ij} \neq a_{ji}$ e $b_{ij} \neq b_{ji}$.

O termo aleatório utilizado nesta equação representa a premissa tomada na elaboração do modelo. O modelo NRTL é tido como um modelo de composição local, tal qual os modelos de Wilson, UNIQUAC e UNIFAC (TERRON, 2009). Ele considera que a distribuição de concentrações ao redor de uma molécula não é totalmente aleatória, sendo diferente, portanto, da concentração global da solução. Isto se deve às diferentes interações presentes entre as moléculas do sistema (PRAUSNITZ *et al.*, 1999).

O modelo NRTL é vantajoso quando se estudam sistemas que se distanciam da idealidade, especialmente para sistemas parcialmente imiscíveis.

3.2 APRENDIZADO DE MÁQUINA

Os modelos baseados em aprendizado de máquina visam utilizar os dados de um problema para buscar padrões e prever os resultados possíveis com base nestes padrões. Neste caso, há por trás de um modelo um conjunto complexo de equações matemáticas que será utilizado pela máquina para que se obtenha um resultado plausível.

Classificam-se os modelos de aprendizado de máquina em dois grupos principais: os modelos supervisionados e os não-supervisionados. Os modelos supervisionados recebem

como entrada os dados e as respostas esperadas para estes dados, de modo que o algoritmo possa identificar os padrões presentes nos dados e criar um modelo que resulte nas respostas fornecidas, enquanto os não-supervisionados recebem somente os dados, devendo utilizar os padrões para gerar uma resposta que não o foi fornecida (IZBICKI, 2020).

Nos modelos supervisionados, o ajuste é realizado por meio de hiperparâmetros que variam para adaptá-lo ao problema estudado. De maneira diferente dos parâmetros de interação binária do modelo NRTL, não necessariamente uma variação nos hiperparâmetros resultará em um melhor ou pior ajuste. Existem algoritmos que automatizam a busca pelo conjunto de hiperparâmetros ótimos, porém eles demandam um grande esforço computacional e tornam o trabalho complexo e demorado. Uma alternativa que pode ser empregada é a pesquisa manual pelos hiperparâmetros ótimos, testando caso a caso.

3.2.1 Considerações matemáticas e de análise

Em qualquer modelo, a escolha dos parâmetros e das variáveis é crucial por diversos motivos. No caso dos modelos de aprendizado de máquina, existem dois fatores que podem tornar a predição ruim: a dimensionalidade e o ajuste fino dos parâmetros, bem como a quantidade destes (HARRELL, 2001).

Quando se utilizam parâmetros em número excessivo em um modelo, o resultado pode ser bom para os dados utilizados na regressão, mas ruim para dados novos. Quando isto ocorre, diz-se que o modelo está superajustado. Este problema é especialmente ruim para os modelos de aprendizado de máquina quando comparados com o modelo NRTL. No modelo NRTL há apenas um conjunto de parâmetros ótimos para ser encontrado que é próprio do modelo e corresponde à mistura estudada, enquanto nos modelos de aprendizado de máquina isto não ocorre, pois aqui não levamos em conta a natureza do problema, somente os padrões existentes nos dados.

Para evitar o superajuste ou detectá-lo sem que seja necessária a busca por novos dados, lança-se mão da estratégia de divisão do conjunto de dados entre treinamento e teste. Para evitar a divisão dos dados em dois conjuntos com padrões semelhantes, o que poderia mascarar o problema do superajuste, este corte é feito de maneira aleatória pelo próprio algoritmo. Desta forma, treina-se o modelo com dados escolhidos aleatoriamente com um determinado conjunto de hiperparâmetros. Se o resultado for satisfatório, realiza-se o teste do modelo treinado com os dados separados para isto e se observa o resultado. Se o ajuste for ruim, o modelo se encontra superajustado aos dados de treinamento, logo se realiza uma nova

divisão em novos conjuntos de treinamento e teste e se procede novamente com a regressão até que tanto o treinamento quanto o teste obtenham resultados bons (IZBICKI, 2020).

Outro problema se relaciona à dimensionalidade do problema. Muitas variáveis produzem um amplo campo de pesquisa de hiperparâmetros, tornando difícil o encontro do conjunto ótimo desejado. É desejável sempre se priorizar o menor número de variáveis possível para que o modelo possa ser, de uma maneira geral, preditivo sem depender de muitos fatores do problema, correndo-se o risco de torná-lo específico para o problema em estudo (IZBICKI, 2020).

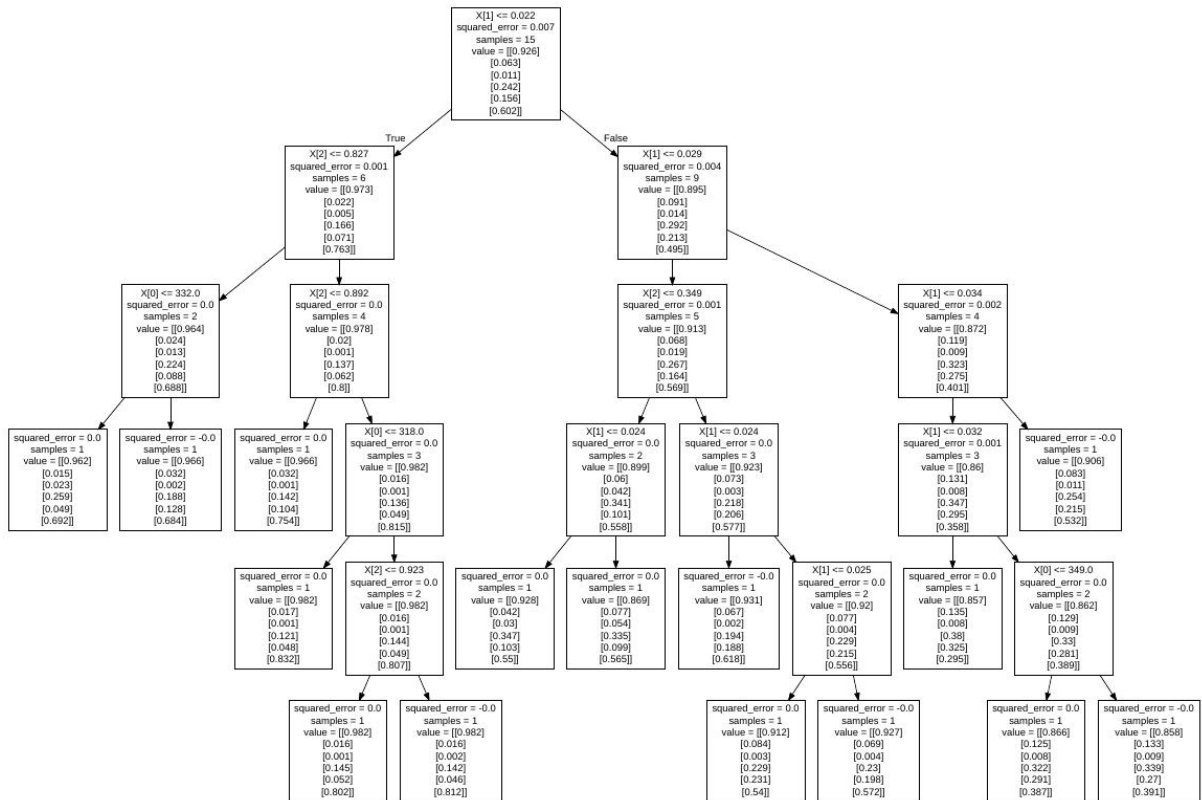
Para se avaliar a qualidade do ajuste, utilizar-se-á o coeficiente de determinação (R^2) ajustado, baseado no coeficiente fornecido pelo próprio modelo após a regressão. Em geral, o coeficiente do treinamento será sempre próximo de 1, sendo válido para o estudo da qualidade do modelo, portanto, o coeficiente do teste.

3.2.2 Florestas aleatórias

O modelo de florestas aleatórias é um modelo baseado em ensemble learning, que consiste em embutir diversos modelos simples que, trabalhando juntos, produzem um resultado confiável de maneira robusta.

Neste modelo são geradas múltiplas árvores de decisão diferentes que receberão como entrada o mesmo conjunto de dados e que produzirão um resultado combinado por meio de uma métrica a ser definida no treinamento, geralmente uma média ponderada. Desta forma, trabalhando de formas diferentes os mesmos dados por meio de árvores diferentes, o problema do superajuste é evitado (IZBICKI, 2020).

Figura 1 – Transcrição de uma árvore de decisão utilizada neste trabalho



Fonte: De autoria própria.

3.3 PROGRAMA UTILIZADO

Neste trabalho, os dados serão analisados por um programa escrito em Python. Esta linguagem de programação apresenta uma grande adaptabilidade no estudo do equilíbrio de fases e no aprendizado de máquina, possuindo pacotes próprios para tal, além de possuir uma escrita fluida e de fácil leitura. Além disso, o Python possui uma vantagem interessante sobre outras linguagens científicas mais bem estabelecidas: Ele foi criado para ser uma interface *open-source*, ou seja, de distribuição e utilização gratuitas. Isto faz com que esta linguagem possua uma vasta comunidade de utilizadores, o que facilita a resolução de problemas de ordem computacional, além de permitir o desenvolvimento de cada vez mais pacotes aplicáveis em diversas áreas do conhecimento.

Além do ponto citado, escrever um programa apresenta como vantagem sobre simuladores a possibilidade de aperfeiçoamento dos mecanismos utilizados na resolução de problemas, bem como que seja feita a avaliação da influência de diversos fatores sobre o estudo de maneira rápida e simples.

4 METODOLOGIA

Foram coletados dados da literatura para o sistema água + ácido acético + acetato para ajuste dos modelos termodinâmico e de aprendizado de máquina, de modo que se viabilize a aplicação destes modelos à dados experimentais.

4.1 ESTUDO TERMODINÂMICO DO EQUILÍBRIO DE FASES

Foi proposto um modelo termodinâmico capaz de modelar os dados experimentais com exatidão e prever dados acerca do equilíbrio existente entre água, ácido acético e acetatos. A análise dos dados seguiu a seguinte rotina:

- Coletaram-se da literatura dados acerca do equilíbrio líquido-líquido existente entre os compostos a serem estudados;
- Executou-se a regressão dos dados experimentais obtidos, de modo a se determinar a performance dos modelos por acetato, por par de ésteres isômeros e com todo o conjunto de dados;
- Realizou-se a predição dos dados do equilíbrio, de modo a se avaliar a qualidade do ajuste realizado pelos modelos.

4.1.1 Equacionamento do problema

A formulação deste problema engloba um conjunto de 6 variáveis: x_1 , x_2 e x_3 , como composições de 1, 2 e 3 na fase aquosa, e y_1 , y_2 e y_3 , como composições de 1, 2 e 3 na fase orgânica.

Para a formulação do problema, desconsiderou-se o fator de correção de Poynting para altas pressões, pois os estudos se desenvolveram na pressão atmosférica. Logo, a equação do equilíbrio de fases (equação 6) para o sistema bifásico se resume à equação 18:

$$(\mathbf{y}_i * \mathbf{P}_i^{\text{SAT}} * \mathbf{x}_i)_{\text{aq}} = (\mathbf{y}_i * \mathbf{P}_i^{\text{SAT}} * \mathbf{y}_i)_{\text{org}} \quad (18)$$

Sendo i o mesmo componente em ambas as fases, considerando que o equilíbrio de fases também é composto pelo equilíbrio térmico e renomeando x_i na fase orgânica como y_i temos a equação 19:

$$(\mathbf{y}_i * \mathbf{x}_i)_{\text{aq}} = (\mathbf{y}_i * \mathbf{y}_i)_{\text{org}} \quad (19)$$

Considerando as fases como soluções ideais, podemos escrever, para cada uma delas as relações presentes nas equações 20 e 21:

$$\mathbf{v}_{\text{aq}} = \sum_i^3 \mathbf{x}_i * \bar{\mathbf{v}}_i \quad (20)$$

$$\mathbf{v}_{\text{org}} = \sum_i^3 \mathbf{y}_i * \bar{\mathbf{v}}_i \quad (21)$$

Onde $\bar{\mathbf{v}}_i$ é o volume molar parcial do composto i e \mathbf{v}_{aq} e \mathbf{v}_{org} são os volumes molares da mistura nas fases aquosa e orgânica. Porém, é mais prático experimentalmente trabalhar com a densidade destas fases, por serem propriedades mais fáceis de serem medidas. Desta forma, considerando as conversões para cada uma das fases nas equações 22 e 23:

$$\bar{\mathbf{v}}_i = \frac{\mathbf{x}_i * \mathbf{M}_i}{\mathbf{M}_{\text{aq}} * \rho_i} \quad (22)$$

$$\bar{\mathbf{v}}_i = \frac{\mathbf{y}_i * \mathbf{M}_i}{\mathbf{M}_{\text{org}} * \rho_i} \quad (23)$$

sendo \mathbf{M}_i a massa molar do composto i , \mathbf{M}_{org} e \mathbf{M}_{aq} as massas molares médias das misturas das fases e ρ_i a densidade de i , e o cálculo da massa molar média nas equações 24 e 25:

$$\mathbf{M}_{\text{aq}} = \sum_i^3 \mathbf{x}_i * \mathbf{M}_i \quad (24)$$

$$\mathbf{M}_{\text{org}} = \sum_i^3 \mathbf{y}_i * \mathbf{M}_i \quad (25)$$

tem-se as equações 26 e 27, correspondentes às equações 21 e 22 em termos de densidade:

$$\frac{1}{\rho_{\text{aq}}} = \sum_i^3 \frac{\mathbf{x}_i * \mathbf{M}_i}{\mathbf{M}_{\text{aq}} * \rho_i} \quad (26)$$

$$\frac{1}{\rho_{\text{org}}} = \sum_i^3 \frac{\mathbf{y}_i * \mathbf{M}_i}{\mathbf{M}_{\text{org}} * \rho_i} \quad (27)$$

O modelo de aprendizado de máquina foi treinado fornecendo os valores de densidade das fases como dados de entrada e as composições das fases como saída.

4.2 CARACTERÍSTICAS FÍSICAS DOS DADOS E INDICADORES DE QUALIDADE

Todos os modelos utilizados dependem de informações como a massa molar das substâncias e as densidades nas diversas temperaturas estudadas. Estes dados constam na Tabela 1. Por conta da faixa restrita de temperaturas, todos os componentes das misturas (acetatos, água e ácido acético) experimentam variação desprezível na densidade. Por conta disso, foram adotadas para o estudo as densidades dos componentes nas condições ambientes de temperatura e pressão (300 K e 1 atm).

Tabela 1 – Massas molares e densidades nas condições ambientes dos componentes das misturas estudadas

M (g/mol)	ρ (g/mL)	Composto
18,015	0,9940	Água
60,053	1,0400	Ácido acético
130,187	0,8293	Acetato de amila
130,187	0,8369	Acetato de isoamila
116,158	0,8760	Acetato de butila
116,158	0,8495	Acetato de isobutila

Fonte: De autoria própria.

Para validação do modelo, utilizaram-se dados relativos a reação entre etanol e ácido acético em excesso, produzindo éster e água. As características dos dados experimentais constam nas Tabelas 2 e 3, onde n é o número de dados por temperatura.

Tabela 2 – Características físicas dos dados experimentais relativos à fase aquosa para o sistema água (1) + ácido acético (2) + éster (3) em equilíbrio

n	T (K)	Δx_1	Δx_2	Δx_3	Éster
4	298,15	0,8689 – 0,9621	0,0150 – 0,0771	0,0229 – 0,0539	Acetato de amila ^[1]
9	304,15	0,8568 – 0,9995	0,0000 – 0,1353	0,0005 – 0,0079	
8	332,15	0,8046 – 0,9992	0,0000 – 0,1779	0,0008 – 0,0175	
8	366,15	0,8027 – 0,9991	0,0000 – 0,1720	0,0009 – 0,0253	
6	293,00	0,8066 – 0,9670	0,0327 – 0,1700	0,0003 – 0,0234	Acetato de butila ^{[3][4]}
5	313,00	0,8424 – 0,9684	0,0312 – 0,1431	0,0004 – 0,0145	Acetato de isobutila ^{[2][3]}
5	293,00	0,8075 – 0,9350	0,0542 – 0,1686	0,0010 – 0,0373	
17	298,00	0,7756 – 0,9994	0,0000 – 0,1970	0,0004 – 0,0275	
5	313,00	0,8013 – 0,9346	0,0636 – 0,1714	0,0018 – 0,0273	Acetato de isoamila ^[3]
5	293,00	0,8122 – 0,9118	0,0872 – 0,1747	0,0010 – 0,0131	
5	313,00	0,8024 – 0,9270	0,0728 – 0,1833	0,0002 – 0,0143	

Fonte: [1] WANG *et al.*, 2007; [2] PROCHAZKA *et al.*, 1996; [3] ROQUE *et al.*, 2018; [4] INCE *et al.*, 2002

Para avaliar a qualidade da regressão, fez-se o cálculo do coeficiente de determinação ajustado (R^2 ajustado) (equação 28), que é uma correção do coeficiente de determinação que visa penalizar modelos em função do seu número de parâmetros ajustáveis e do tamanho do conjunto de dados utilizado.

$$\mathbf{R}^2_{\text{ajustado}} = \mathbf{1} - \left(\frac{(\mathbf{1} - \mathbf{R}^2) * (\mathbf{n} - \mathbf{1})}{(\mathbf{n} - \mathbf{k} - \mathbf{1})} \right) \quad (28)$$

Onde R^2 é o coeficiente de determinação convencional, n é o tamanho do conjunto de dados e k é o número de parâmetros ajustáveis. Sabe-se que melhor será a qualidade da regressão quanto mais próximo de 1 for o coeficiente de determinação ajustado.

Calculou-se, também, o conjunto de resíduos (r) entre as composições experimental e calculada para cada componente i em cada fase. Para a fase aquosa (1), tem-se a equação 29 e para a fase orgânica (2), a equação 30:

$$\mathbf{r}_{i,1} = \mathbf{x}_{i,\text{calc}} - \mathbf{x}_{i,\text{exp}} \quad (29)$$

$$\mathbf{r}_{i,2} = \mathbf{y}_{i,\text{calc}} - \mathbf{y}_{i,\text{exp}} \quad (30)$$

Um bom modelo produz resíduos próximos de uma distribuição normal e não correlacionados. Desta forma, uma boa avaliação da qualidade do modelo deve passar pelo estudo dos resíduos.

Tabela 3 – Características físicas dos dados experimentais relativos à fase orgânica para o sistema água (1) + ácido acético (2) + éster (3) em equilíbrio

n	T (K)	Δy_1	Δy_2	Δy_3	Fonte
4	298,15	0,2588 - 0,3468	0,0491 - 0,1032	0,5500 - 0,6921	Acetato de amila ^[1]
9	304,15	0,0750 - 0,3795	0,0750 - 0,3251	0,2954 - 0,9250	
8	332,15	0,1055 - 0,3857	0,1055 - 0,3106	0,3037 - 0,8945	
8	366,15	0,1126 - 0,4081	0,1126 - 0,3017	0,2902 - 0,8874	
6	293,00	0,1189 - 0,4475	0,0680 - 0,2879	0,2646 - 0,8131	Acetato de butila ^{[3][4]}
5	313,00	0,1414 - 0,3861	0,0695 - 0,2553	0,3586 - 0,7891	
5	293,00	0,1570 - 0,4287	0,1312 - 0,2947	0,2766 - 0,7118	Acetato de isobutila ^{[2][3]}
17	298,00	0,0816 - 0,4601	0,0007 - 0,2870	0,2529 - 0,9177	
5	313,00	0,1869 - 0,4248	0,1255 - 0,2757	0,2995 - 0,6876	
5	293,00	0,1116 - 0,3029	0,1137 - 0,3025	0,3946 - 0,7747	Acetato de isoamila ^[3]
5	313,00	0,1607 - 0,3268	0,1220 - 0,2849	0,3883 - 0,7173	

Fonte: [1] WANG *et al.*, 2007; [2] PROCHAZKA *et al.*, 1996; [3] ROQUE *et al.*, 2018; [4] INCE *et al.*, 2002

4.3 TESTE DE CORRELAÇÃO DE OTHMER-TOBIAS

A equação de Gibbs-Duhem é largamente utilizada no teste da consistência termodinâmica dos dados de equilíbrio ao correlacionar grandezas molares parciais e frações dos componentes entre as fases do sistema. Porém, sua aplicação encontra problemas ao tentar avaliar dados de equilíbrio líquido-líquido (CARNITI *et al.*, 1978).

Por exemplo, há no equilíbrio líquido-vapor somente uma fase líquida, o que simplifica o cálculo do coeficiente de atividade e, por consequência, a aplicação da equação de Gibbs-Duhem. Porém, no equilíbrio líquido-líquido, por coexistirem duas fases líquidas, a determinação do coeficiente de atividade é mais complexa. Neste caso, não se discute consistência dos dados no equilíbrio líquido-líquido, mas correlação, a qual é avaliada por modelos semi-empíricos, como o proposto por Othmer e Tobias (1942) que se encontra representado pela equação 31. A correlação entre os dados experimentais indica a qualidade destes e a validade de sua aplicação em modelos termodinâmicos.

$$\ln \left(\frac{1 - w_{3,org}}{w_{3,org}} \right) = A + B * \ln \left(\frac{1 - w_{1,aq}}{w_{1,aq}} \right) \quad (31)$$

Onde $w_{3,org}$ e $w_{1,aq}$ são as frações mássicas de éster na fase orgânica e de água na fase aquosa, respectivamente.

Como o teste de correlação de Othmer-Tobias utiliza fração mássica, foi necessária a conversão dos dados em fração molar dos componentes com a equação 32:

$$w_i = \frac{x_i * M}{M_i} \quad (32)$$

onde M é a massa molar média do sistema no ponto a ser convertido e M_i é a massa molar do componente.

Por se tratar de um modelo linear, quanto mais próximo de 1 o coeficiente de determinação estiver, maior será a correlação entre os dados.

5 RESULTADOS E DISCUSSÃO

5.1 AVALIAÇÃO DOS DADOS EXPERIMENTAIS

A avaliação da correlação entre os dados experimentais foi feita utilizando o teste de Othmer-Tobias (equação 31). Os coeficientes de determinação para cada éster e por temperatura se encontram na Tabela 4.

Tabela 4 – Coeficientes de determinação do teste de Othmer-Tobias por temperatura para cada éster estudado

Éster	T (K)	R ²
Acetato de amila	298,00	0,5565
	304,00	0,9810
	332,00	0,9943
	366,00	0,9782
Acetato de isoamila	293,00	0,9980
	313,00	0,9943
Acetato de butila	293,00	0,8795
	313,00	0,9891
Acetato de isobutila	293,00	0,9203
	298,00	0,9972
	313,00	0,9313

Fonte: De autoria própria.

5.2 RESULTADOS DOS DIFERENTES MODELOS DE REGRESSÃO

A Tabela 5 dispõe os parâmetros de interação binária para o modelo NRTL obtidos via regressão para os seguintes casos:

- Considerando cada éster;
- Avaliando os pares de isômeros acetato de isoamila-acetato de amila e acetato de isobutila-acetato de butila.

As regressões por métodos de aprendizado de máquina foram feitas fixando os mesmos hiperparâmetros para todos os conjuntos de dados. As florestas aleatórias foram construídas considerando uma floresta com 10 árvores e o treinamento deste modelo foi feito com 60% dos dados.

Nos tópicos abaixo, serão discutidos os resultados obtidos pelos três métodos de regressão para cada conjunto de dados.

Tabela 5 – Parâmetros de ajuste obtidos pela regressão do modelo NRTL por conjunto de dados

Conjunto de dados	Acetato de amila	Acetato de isobutila
a ₁₂	74,5895	1,0000
a ₂₁	-79,8685	1,0000
a ₁₃	-431,8471	1,0000
a ₃₁	434,2099	1,0000
a ₂₃	-260,9952	1,0000
a ₃₂	250,9932	1,0000
b ₁₂	-64440,9732	1,0000
b ₂₁	66280,5292	1,0000
b ₁₃	-48840,7410	1,0000
b ₃₁	48793,4052	1,0000
b ₂₃	-48533,8704	1,0000
b ₃₂	51005,2106	1,0000

Fonte: De autoria própria.

5.2.1 Acetato de amila

5.2.1.1 Modelo NRTL

A Tabela 6 contém os coeficientes de determinação dos dados para a mistura contendo o éster acetato de amila por componente e por fase.

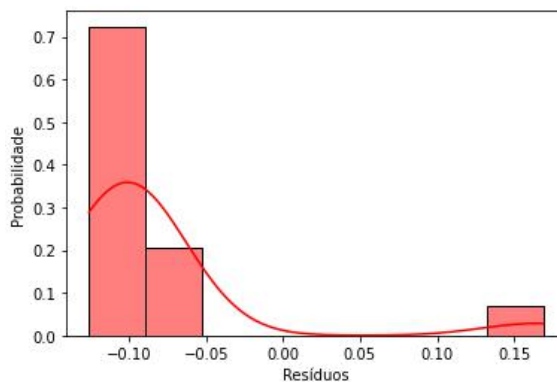
Tabela 6 – Coeficientes de determinação ajustados por fase e por componente da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL

Fase	Componente	R ² ajustado
Aquosa	Água	0,9893
	Ácido acético	0,4281
	Éster	0,2877
Orgânica	Água	0,8684
	Ácido acético	0,4211
	Éster	0,9833

Fonte: De autoria própria.

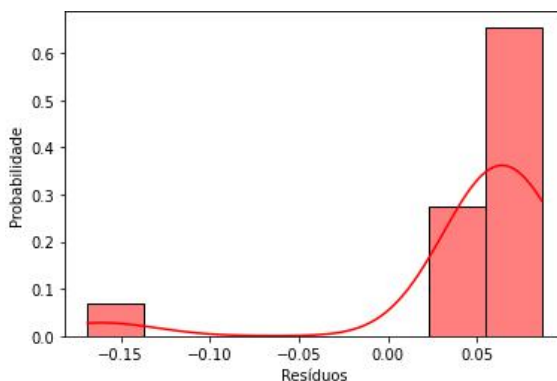
As figuras de 2 a 7 mostram a distribuição dos resíduos da regressão por componente e por fase para o sistema água + ácido acético + acetato de amila.

Figura 2 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL



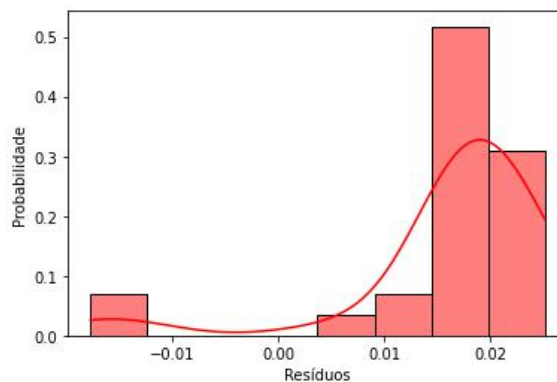
Fonte: De autoria própria.

Figura 3 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL



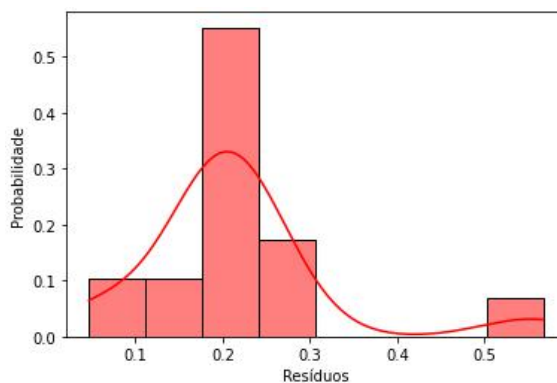
Fonte: De autoria própria.

Figura 4 - Distribuição dos resíduos das frações de acetato de amila na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL



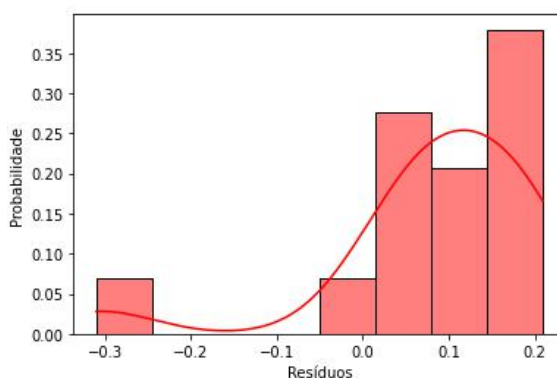
Fonte: De autoria própria.

Figura 5 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL



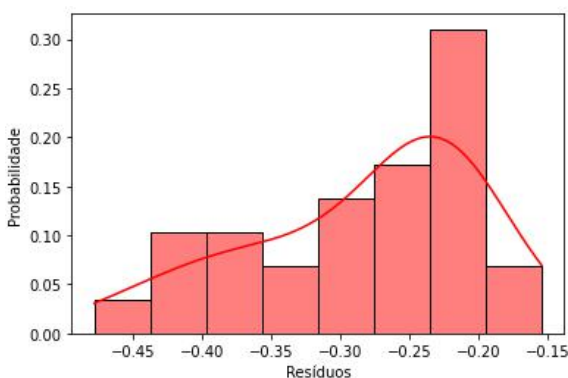
Fonte: De autoria própria.

Figura 6 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL



Fonte: De autoria própria.

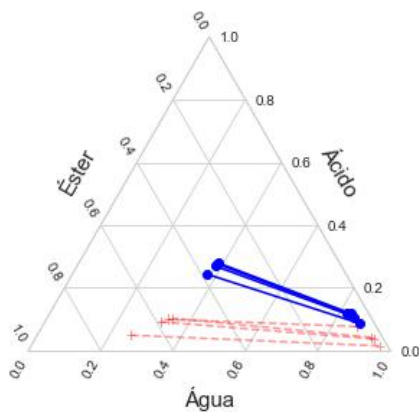
Figura 7 - Distribuição dos resíduos das frações de acetato de amila na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo NRTL



Fonte: De autoria própria.

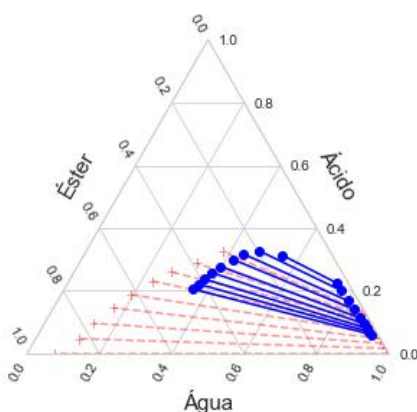
As figuras de 8 a 11 mostram a os diagramas ternários por temperatura para o sistema água + ácido acético + acetato de amila.

Figura 8 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 298,00K (+ = Experimental, o = Calculado utilizando o modelo NRTL)



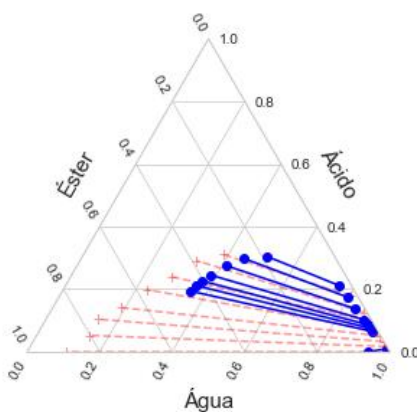
Fonte: De autoria própria.

Figura 9 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 304,00K (+ = Experimental, o = Calculado utilizando o modelo NRTL)



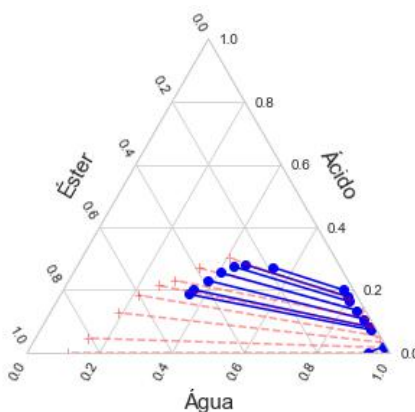
Fonte: De autoria própria.

Figura 10 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 332,00K (+ = Experimental, o = Calculado utilizando o modelo NRTL)



Fonte: De autoria própria.

Figura 11 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 366,00K (+ = Experimental, o = Calculado utilizando o modelo NRTL)



Fonte: De autoria própria.

5.2.1.2 Modelo de florestas aleatórias

A Tabela 7 contém os coeficientes de determinação dos dados para o acetato de amila por componente e por fase.

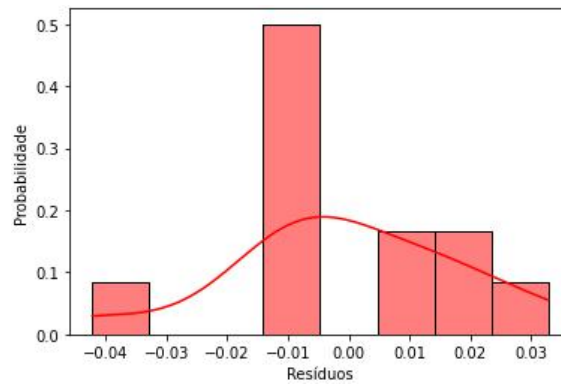
Tabela 7 – Coeficientes de determinação por fase da regressão e por componente para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias

Fase	Componente	R ² ajustado
Aquosa	Água	0,9993
	Ácido acético	0,9517
	Éster	0,6602
Orgânica	Água	0,9659
	Ácido acético	0,9796
	Éster	0,9903

Fonte: De autoria própria.

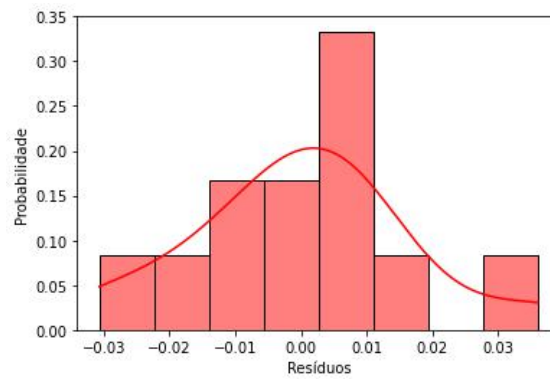
As figuras de 12 a 17 mostram a distribuição dos resíduos da regressão por componente e por fase pra o sistema água + ácido acético + acetato de amila.

Figura 12 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias



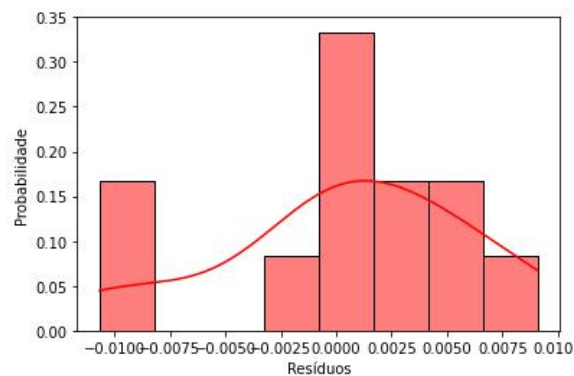
Fonte: De autoria própria.

Figura 13 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias



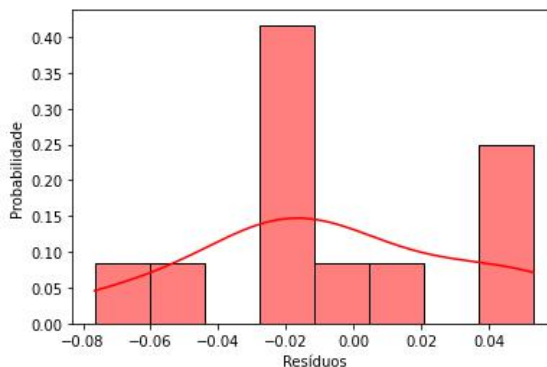
Fonte: De autoria própria.

Figura 14 - Distribuição dos resíduos das frações de acetato de amila na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias



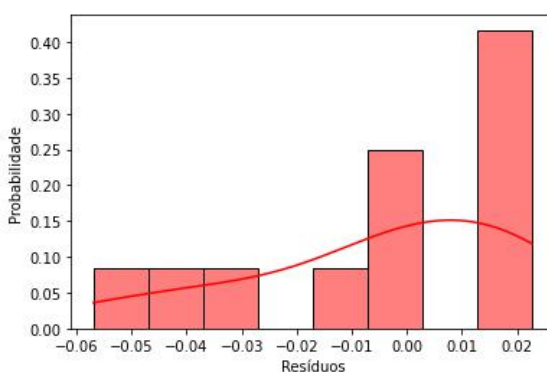
Fonte: De autoria própria.

Figura 15 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias



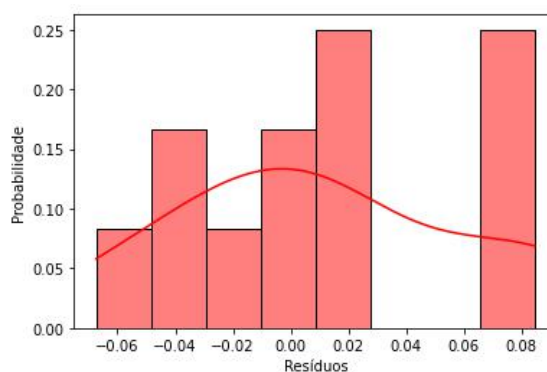
Fonte: De autoria própria.

Figura 16 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias



Fonte: De autoria própria.

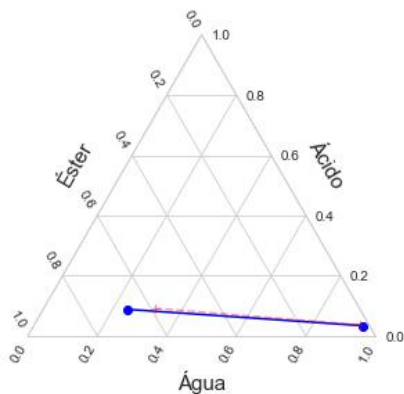
Figura 17 - Distribuição dos resíduos das frações de acetato de amila na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila utilizando o modelo de florestas aleatórias



Fonte: De autoria própria.

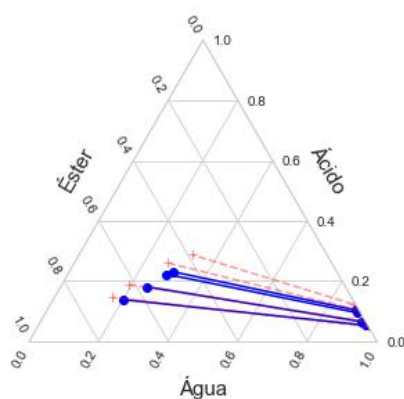
As figuras de 18 a 21 mostram a os diagramas ternários por temperatura para o sistema água + ácido acético + acetato de amila.

Figura 18 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 298,00K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



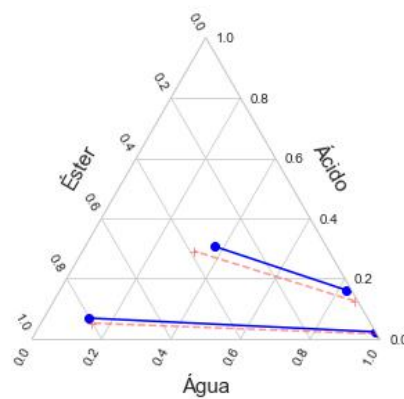
Fonte: De autoria própria.

Figura 19 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 304,00K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



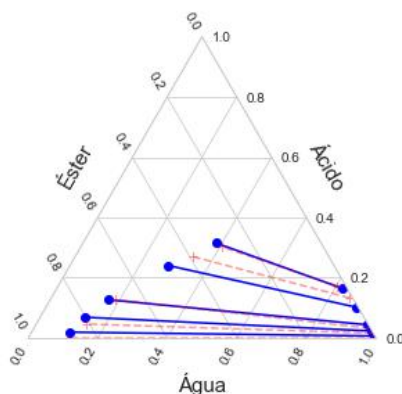
Fonte: De autoria própria.

Figura 20 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 332,00K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



Fonte: De autoria própria.

Figura 21 – Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila na temperatura de 366,00K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



Fonte: De autoria própria.

5.2.2 Acetato de isobutila

5.2.2.1 Modelo NRTL

A Tabela 8 contém os coeficientes de determinação dos dados para o acetato de isobutila por componente e por fase.

Tabela 8 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo NRTL

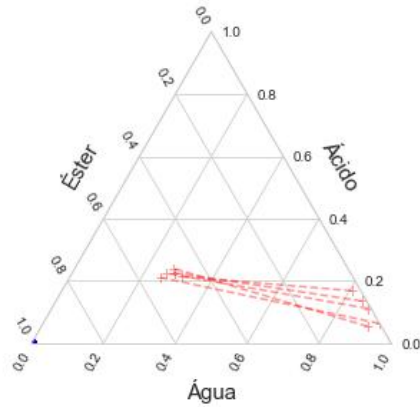
Fase	Componente	R ² ajustado
Aquosa	Água	0,0000
	Ácido acético	0,0000
	Éster	0,0000
Orgânica	Água	0,0000
	Ácido acético	0,0000
	Éster	0,6613

Fonte: De autoria própria.

Assim como no modelo NRTL utilizando o sistema que continha acetato de isoamila, para este caso serão omitidos os diagramas de distribuição dos resíduos da regressão, pois somente os coeficientes de determinação já são suficientes para atestar a qualidade do modelo.

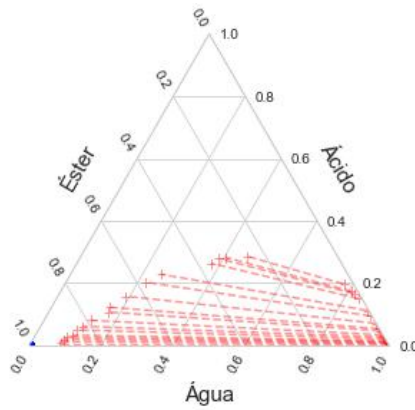
As figuras de 22 a 24 mostram a os diagramas ternários por temperatura para o sistema água + ácido acético + acetato de isobutila.

Figura 22 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 293,00 K (+ = Experimental, o = Calculado utilizando o modelo NRTL)



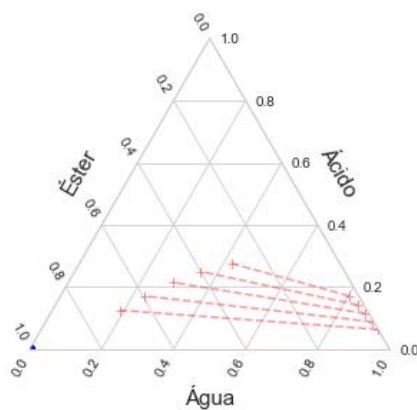
Fonte: De autoria própria.

Figura 23 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 298,00 K (+ = Experimental, o = Calculado utilizando o modelo NRTL)



Fonte: De autoria própria.

Figura 24 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 313,00 K (+ = Experimental, o = Calculado utilizando o modelo NRTL)



Fonte: De autoria própria.

5.2.2.2 Modelo de florestas aleatórias

A Tabela 9 contém os coeficientes de determinação dos dados para o acetato de isobutila por componente e por fase.

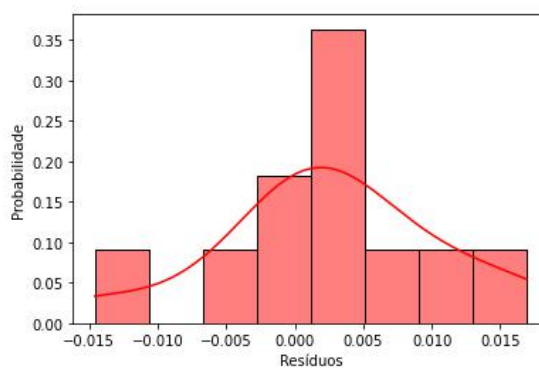
Tabela 9 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias.

Fase	Componente	R ² ajustado
Aquosa	Água	1,0000
	Ácido acético	0,9028
	Éster	0,0000
Orgânica	Água	0,5707
	Ácido acético	0,9721
	Éster	0,9883

Fonte: De autoria própria.

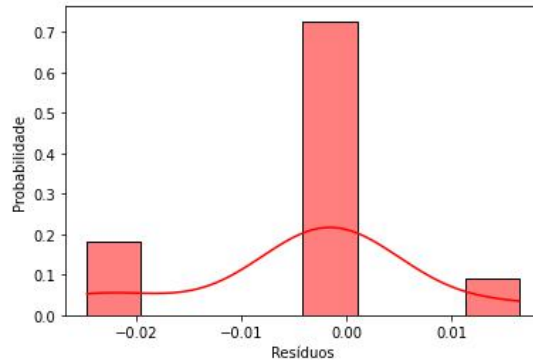
As figuras de 25 a 30 mostram a distribuição dos resíduos da regressão por componente e por fase pra o sistema água + ácido acético + acetato de isobutila.

Figura 25 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias



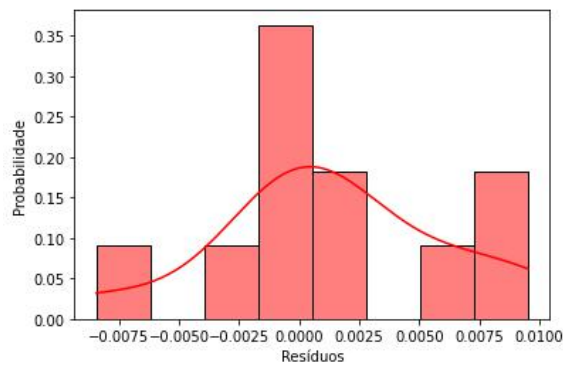
Fonte: De autoria própria.

Figura 26 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias



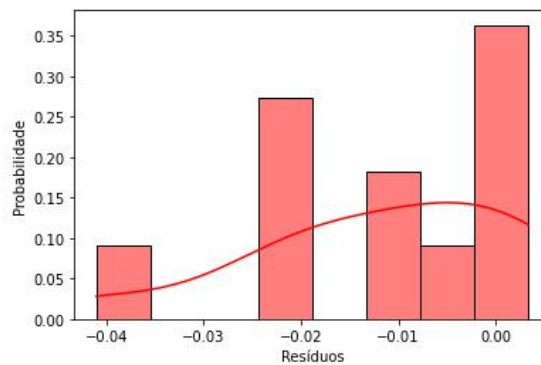
Fonte: De autoria própria.

Figura 27 - Distribuição dos resíduos das frações de acetato de isobutila na fase aquosa da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias



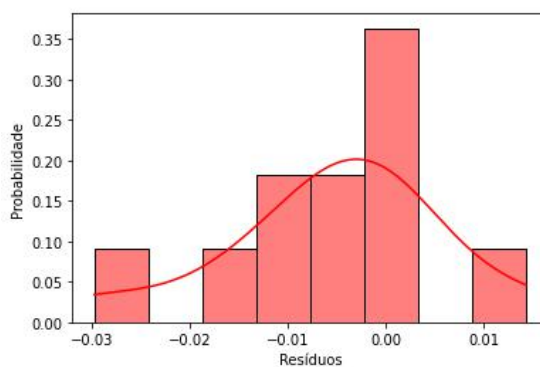
Fonte: De autoria própria.

Figura 28 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias



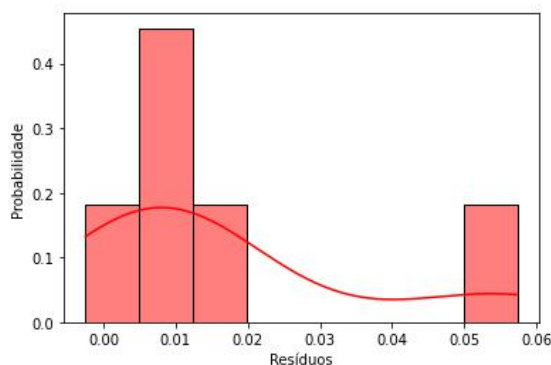
Fonte: De autoria própria.

Figura 29 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias



Fonte: De autoria própria.

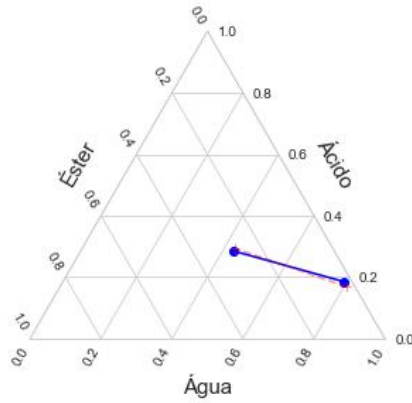
Figura 30 - Distribuição dos resíduos das frações de acetato de isobutila na fase orgânica da regressão para o sistema água + ácido acético + acetato de isobutila utilizando o modelo de florestas aleatórias



Fonte: De autoria própria.

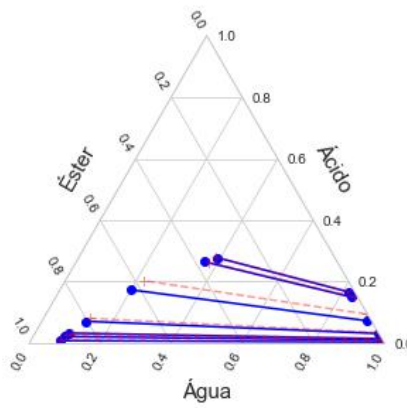
As figuras de 31 a 33 mostram a os diagramas ternários por temperatura para o sistema água + ácido acético + acetato de isobutila.

Figura 31 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 293,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



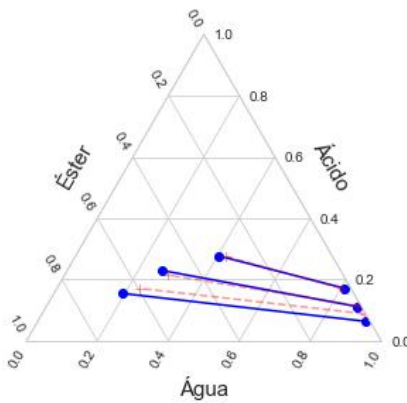
Fonte: De autoria própria.

Figura 32 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 298,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



Fonte: De autoria própria.

Figura 33 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de isobutila na temperatura de 313,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



Fonte: De autoria própria.

5.2.3 Acetato de amila + Acetato de isoamila

Este estudo foi feito considerando o acetato de amila e o acetato de isoamila como um éster somente. Por serem isômeros, a massa molar é a mesma e as densidades são próximas. Porém, a pequena diferença entre as densidades já é suficiente para diminuir a precisão do modelo NRTL, portanto a regressão com este não foi utilizada no estudo.

5.2.3.1 Modelo de florestas aleatórias

A Tabela 10 contém os coeficientes de determinação dos dados para o acetato de isobutila por componente e por fase.

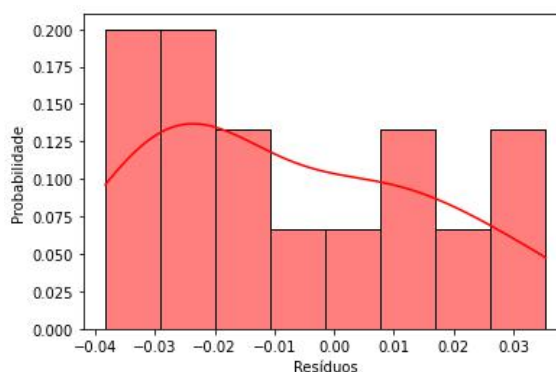
Tabela 10 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias.

Fase	Componente	R ² ajustado
Aquosa	Água	0,9995
	Ácido acético	0,9082
	Éster	0,9793
Orgânica	Água	0,9991
	Ácido acético	0,9287
	Éster	0,9972

Fonte: De autoria própria.

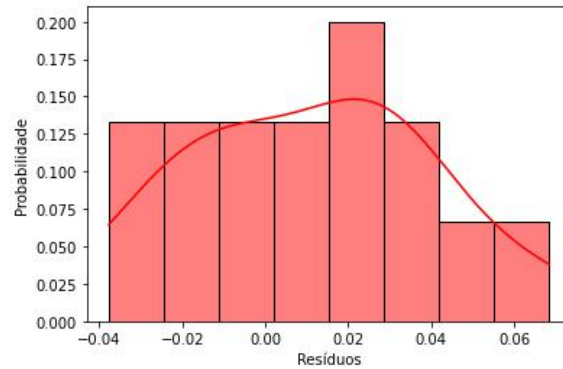
As figuras de 34 a 39 mostram a distribuição dos resíduos da regressão por componente e por fase pra o sistema água + ácido acético + acetato de amila/isoamila.

Figura 34 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias



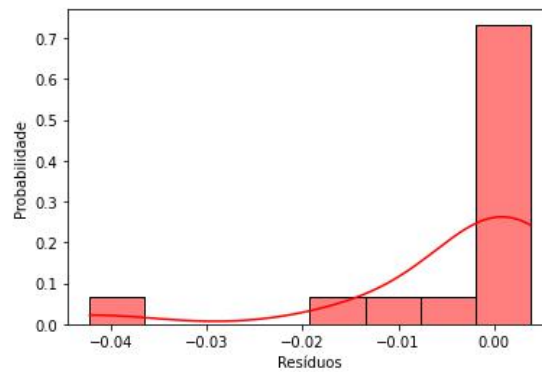
Fonte: De autoria própria.

Figura 35 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias



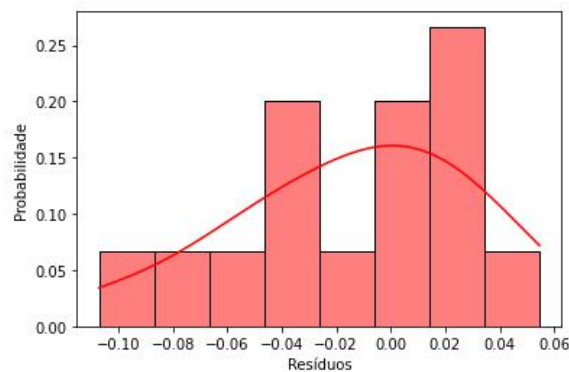
Fonte: De autoria própria.

Figura 36 - Distribuição dos resíduos das frações de acetato de amila/isoamila na fase aquosa da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias



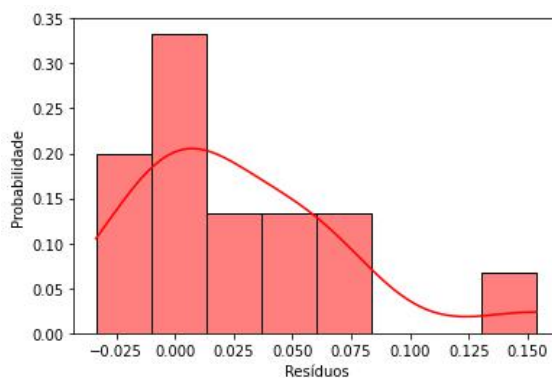
Fonte: De autoria própria.

Figura 37 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias



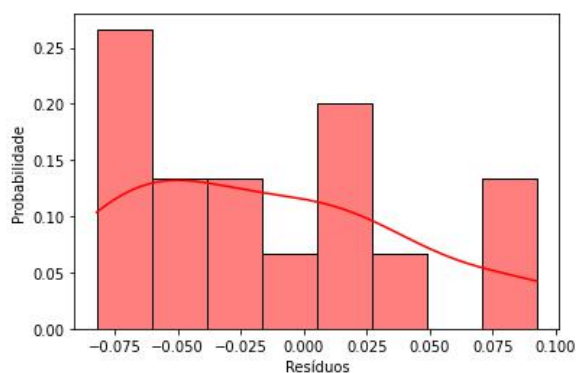
Fonte: De autoria própria.

Figura 38 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias



Fonte: De autoria própria.

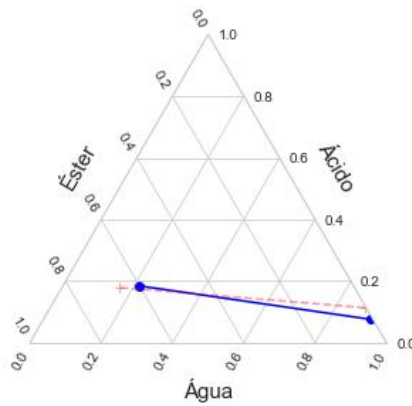
Figura 39 - Distribuição dos resíduos das frações de acetato de amila/isoamila na fase orgânica da regressão para o sistema água + ácido acético + acetato de amila/isoamila utilizando o modelo de florestas aleatórias



Fonte: De autoria própria.

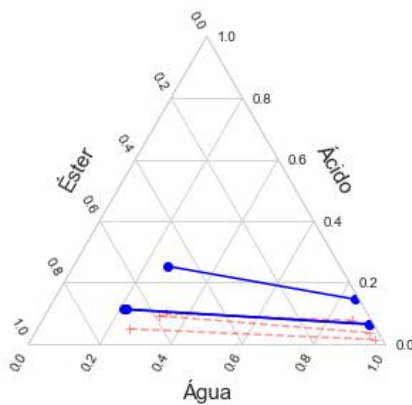
As figuras de 40 a 45 mostram a os diagramas ternários por temperatura para o sistema água + ácido acético + acetato de amila/isoamila.

Figura 40 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 293,00 K(+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



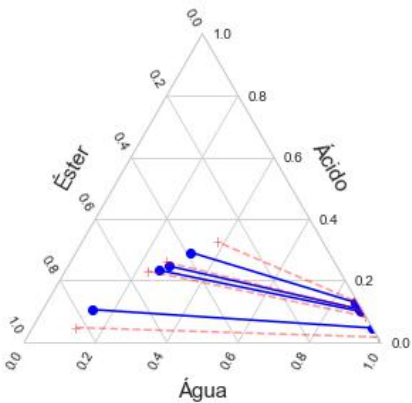
Fonte: De autoria própria.

Figura 41 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 298,00 K(+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



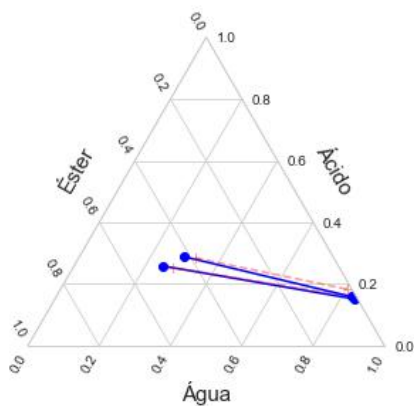
Fonte: De autoria própria.

Figura 42 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 304,00 K(+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



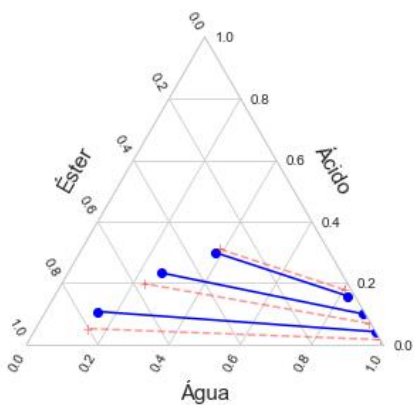
Fonte: De autoria própria.

Figura 43 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 313,00 K(+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



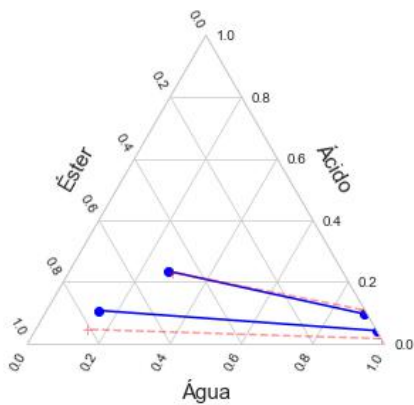
Fonte: De autoria própria.

Figura 44 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 332,00 K(+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



Fonte: De autoria própria.

Figura 45 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de amila/isoamila na temperatura de 366,00 K(+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



Fonte: De autoria própria.

5.2.4 Acetato de butila + Acetato de isobutila

De maneira semelhante ao estudo do sistema anterior, foram considerados um único éster os acetatos de butila e isobutila, aproveitando-se de suas propriedades semelhantes decorrentes da isomeria.

5.2.4.1 Modelo de florestas aleatórias

A Tabela 11 contém os coeficientes de determinação dos dados para o acetato de isobutila por componente e por fase.

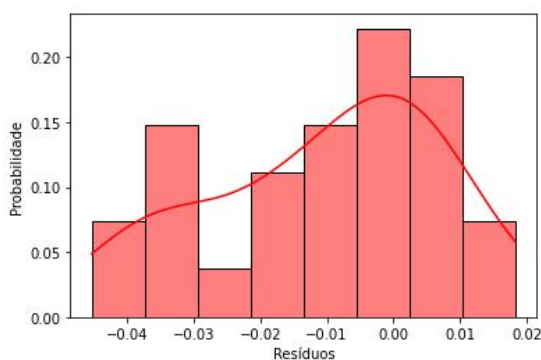
Tabela 11 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetato de butila/isobutila utilizando o modelo de florestas aleatórias.

Fase	Componente	R ² ajustado
Aquosa	Água	0,9998
	Ácido acético	0,9894
	Éster	0,9452
Orgânica	Água	0,9932
	Ácido acético	0,9998
	Éster	0,9998

Fonte: De autoria própria.

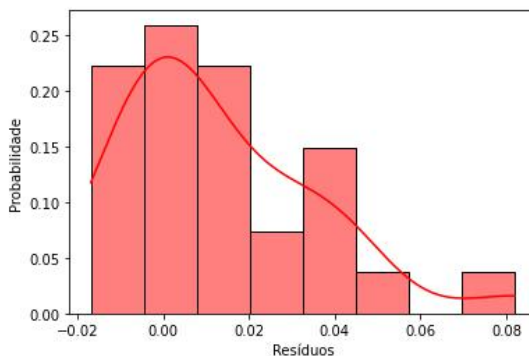
As figuras de 46 a 51 mostram a distribuição dos resíduos da regressão por componente e por fase pra o sistema água + ácido acético + acetato de butila/isobutila.

Figura 46 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetato de butila/isobutila utilizando o modelo de florestas aleatórias



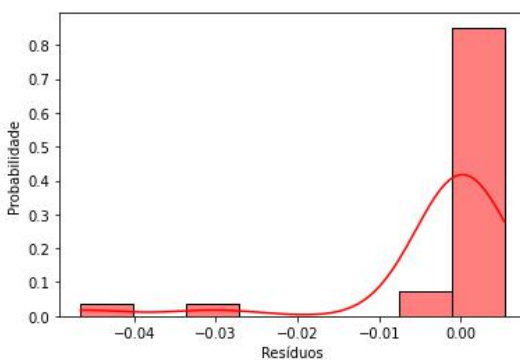
Fonte: De autoria própria.

Figura 47 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetato de butila/isobutila utilizando o modelo de florestas aleatórias



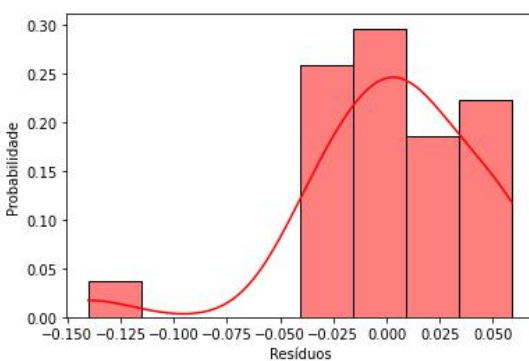
Fonte: De autoria própria.

Figura 48 - Distribuição dos resíduos das frações de acetato de butila/isobutila na fase aquosa da regressão para o sistema água + ácido acético + acetato de butila/isobutila utilizando o modelo de florestas aleatórias



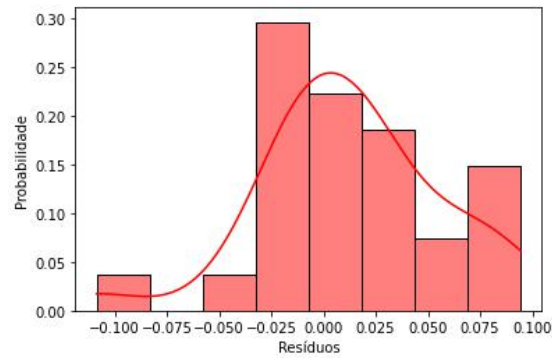
Fonte: De autoria própria.

Figura 49 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetato de butila/isobutila utilizando o modelo de florestas aleatórias



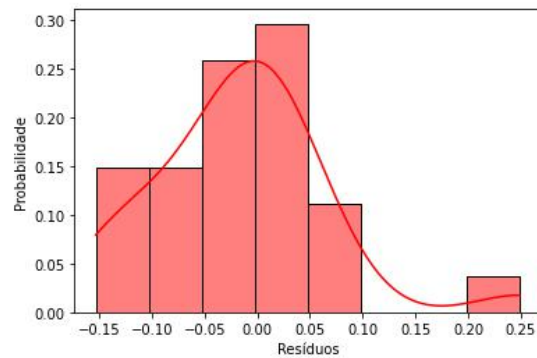
Fonte: De autoria própria.

Figura 50 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetato de butila/isobutila utilizando o modelo de florestas aleatórias



Fonte: De autoria própria.

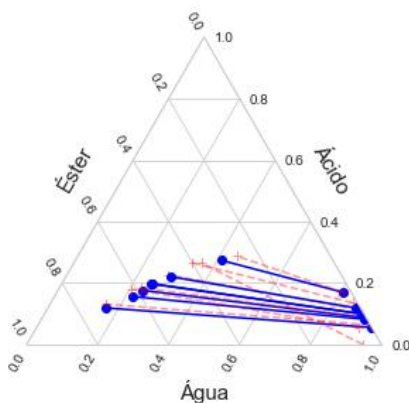
Figura 51 - Distribuição dos resíduos das frações de acetato de butila/isobutila na fase aquosa da regressão para o sistema água + ácido acético + acetato de butila/isobutila utilizando o modelo de florestas aleatórias



Fonte: De autoria própria.

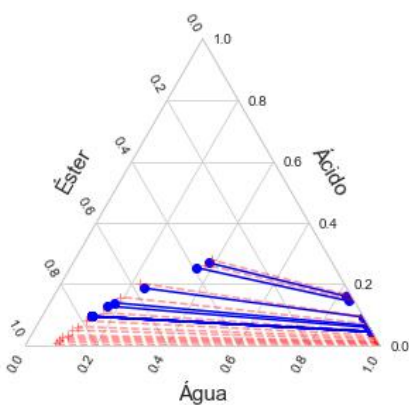
As figuras de 52 a 54 mostram a os diagramas ternários por temperatura para o sistema água + ácido acético + acetato de butila/isobutila.

Figura 52 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de butila/isobutílica na temperatura de 293,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



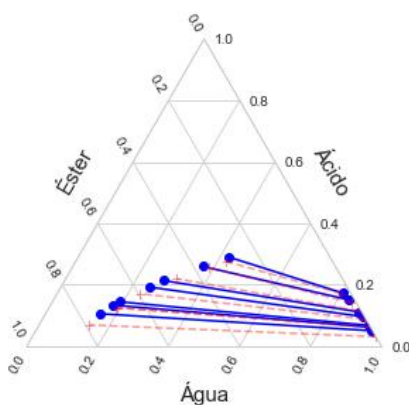
Fonte: De autoria própria.

Figura 53 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de butila/isobutílica na temperatura de 298,00 K(+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



Fonte: De autoria própria.

Figura 54 - Diagrama ternário das composições para o sistema água + ácido acético + acetato de butila/isobutílica na temperatura de 313,00 K(+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



Fonte: De autoria própria.

5.2.5 Todos os ésteres

Este estudo visa simplificar ainda mais a modelagem, considerando todos os ésteres como uma única substância levando em conta que todos possuem a mesma cadeia principal. Da mesma forma que as anteriores, foi realizada a regressão somente por meio do modelo de florestas aleatórias, pois a diferença de massas molares e de densidades poderia prejudicar a performance do modelo NRTL neste caso.

5.2.5.1 Modelo de florestas aleatórias

A Tabela 12 contém os coeficientes de determinação dos dados por componente e por fase.

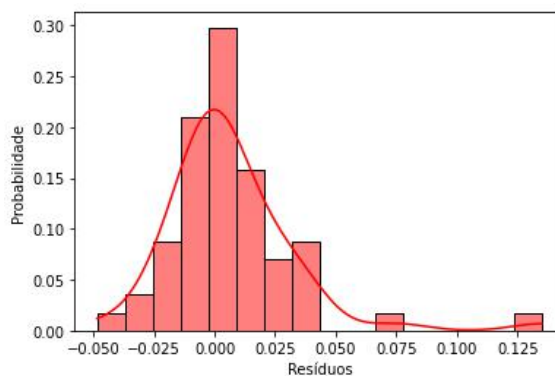
Tabela 12 – Coeficientes de determinação por fase e por componente da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias.

Fase	Componente	R² ajustado
Aquosa	Água	0,9990
	Ácido acético	0,8567
	Éster	0,9519
Orgânica	Água	0,9613
	Ácido acético	0,8304
	Éster	0,9964

Fonte: De autoria própria.

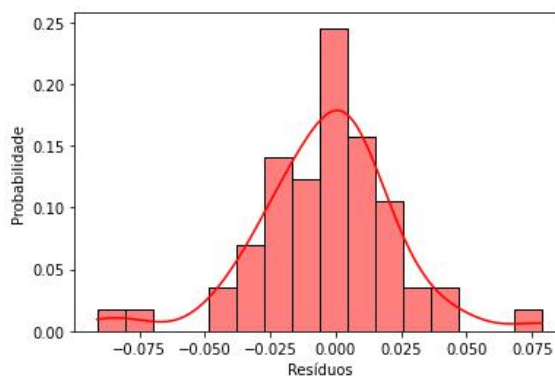
As figuras de 55 a 59 mostram a distribuição dos resíduos da regressão por componente e por fase pra o sistema água + ácido acético + acetatos.

Figura 55 - Distribuição dos resíduos das frações de água na fase aquosa da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias



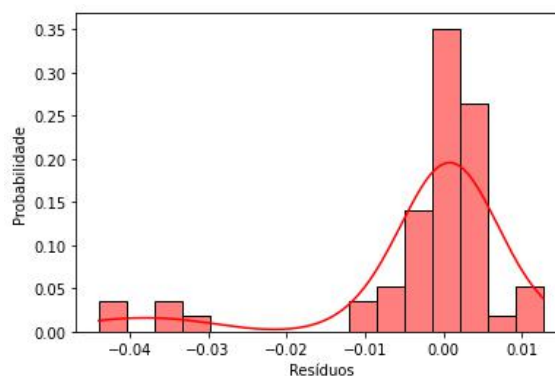
Fonte: De autoria própria.

Figura 56 - Distribuição dos resíduos das frações de ácido acético na fase aquosa da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias



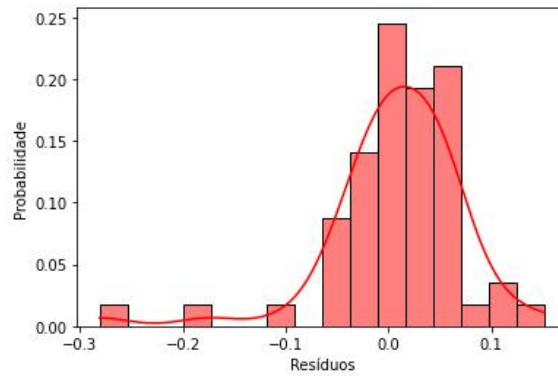
Fonte: De autoria própria.

Figura 57 - Distribuição dos resíduos das frações de acetatos na fase aquosa da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias



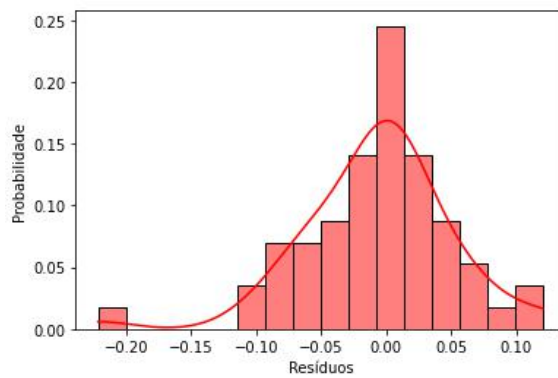
Fonte: De autoria própria.

Figura 58 - Distribuição dos resíduos das frações de água na fase orgânica da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias



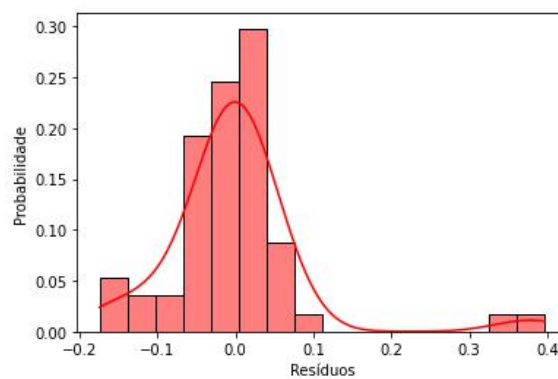
Fonte: De autoria própria.

Figura 59 - Distribuição dos resíduos das frações de ácido acético na fase orgânica da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias



Fonte: De autoria própria.

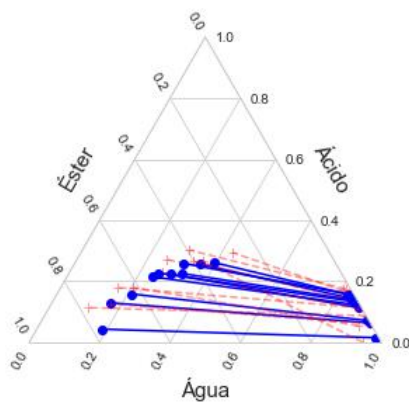
Figura 60 - Distribuição dos resíduos das frações de acetatos na fase orgânica da regressão para o sistema água + ácido acético + acetatos utilizando o modelo de florestas aleatórias



Fonte: De autoria própria.

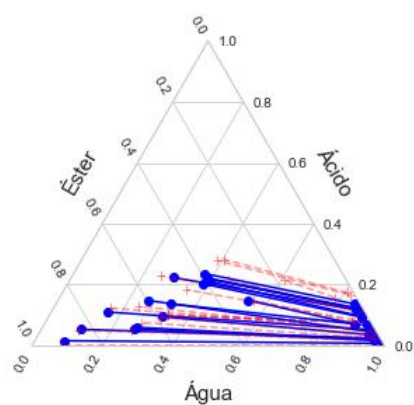
As figuras de 61 a 66 mostram as diagramas ternários por temperatura para o sistema água + ácido acético + acetatos.

Figura 61 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 293,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



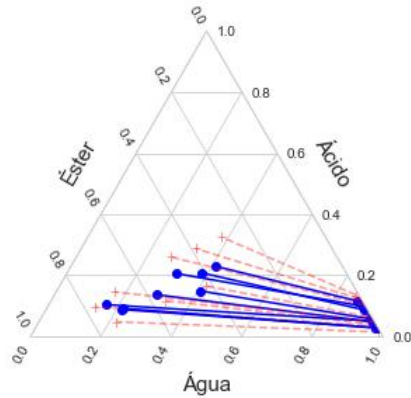
Fonte: De autoria própria.

Figura 62 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 298,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



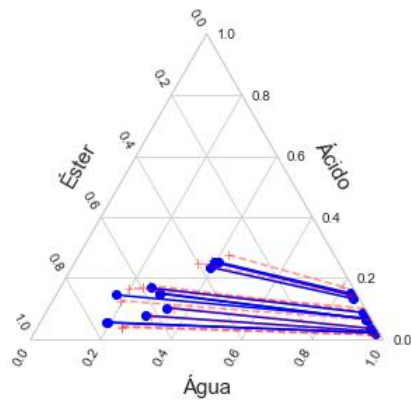
Fonte: De autoria própria.

Figura 63 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 304,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



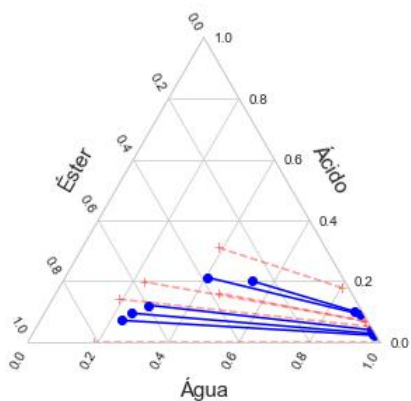
Fonte: De autoria própria.

Figura 64 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 313,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



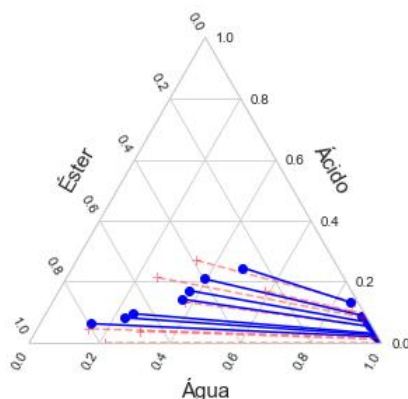
Fonte: De autoria própria.

Figura 65 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 332,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



Fonte: De autoria própria.

Figura 66 - Diagrama ternário das composições para o sistema água + ácido acético + acetatos na temperatura de 366,00 K (+ = Experimental, o = Calculado utilizando o modelo de florestas aleatórias)



Fonte: De autoria própria.

5.3 DISCUSSÃO DOS RESULTADOS

Pode-se notar no exposto no tópico 5.2 resultados variados, desde modelos que ajustaram os dados muito bem até modelos que não ajustaram os dados. Um resultado comum aos dois modelos em todas as análises é a faixa de valores do coeficiente de determinação ajustado. Fica claro que ambos os modelos foram penalizados pelo número de parâmetros elevado (12 no modelo NRTL e 10 no de florestas aleatórias) e pela escassez de dados nas análises individuais.

5.3.1 Modelo NRTL

No caso do modelo NRTL, observa-se que sua performance foi próxima do esperado para o acetato de amila, pois, baseando-se na análise das distribuições dos resíduos para este éster (Figuras de 2 a 7), nota-se uma aproximação de uma normal padrão, mostrando que a grande maioria dos resíduos se avizinhava de 0.

Outra análise que pode ser feita para embasar esta performance é a dos coeficientes de determinação ajustados, que para o acetato de amila indicam uma correlação moderada para forte para praticamente todos os componentes em todas as fases. Observando os diagramas ternários para este éster com este modelo (Figuras de 8 a 11), é visível que há um deslocamento dos pontos calculados em relação ao experimentais.

O desempenho do modelo NRTL ao se ajustar aos dados do acetato de amila pode ser justificado pela qualidade destes. De posse dos coeficientes de determinação obtidos pelo teste de correlação de Othmer-Tobias (Tabela 4), observa-se que à temperatura de 298 K para o acetato de amila, o resultado obtido foi 0,5565, muito aquém do desejado, que é um valor acima de 0,9. Como a regressão é feita utilizando todos os dados disponíveis, não para cada

temperatura, a existência de valores sem correlação forte pode afetar a qualidade do ajuste para todos os demais dados.

O que se observa nos ajustes dos dados para o acetato de isobutila são desvios muito elevados. Pode-se notar que os dados calculados correspondem ao ponto cuja composição de éster é igual a 1 em ambas fases. A qualidade destas regressões pode ser discutida somente considerando os coeficientes de determinação dos componentes nas fases.

Observando os resultados para o acetato de isobutila aplicando-se o modelo NRTL, constata-se a baixa qualidade da regressão por meio dos valores nulos para a correlação dos dados experimentais e calculados do ácido acético em ambas as fases e, também, por conta dos valores do coeficiente de correlação para o éster em ambas as fases, que indicam uma correlação fraca entre os dados experimentais e calculados.

A principal hipótese para todos os erros encontrados até este ponto é o equacionamento do problema. Esta possibilidade se desdobra em dois pontos: quantidade de parâmetros e complexidade do sistema a ser resolvido. Para que o equacionamento deste trabalho pudesse ser empregado na modelagem termodinâmica, foram necessárias simplificações, como a suposição de solução ideal para as fases, que podem ter contribuído para os desvios observados.

Prausnitz (1999) propõe um equacionamento baseado nas composições globais da mistura, que contorna qualquer suposição de idealidade feita no equacionamento deste trabalho. Porém, nas referências que forneceram os dados utilizados neste trabalho, não foram destacados valores de composição global para nenhuma das misturas estudadas.

5.3.2 Modelo de florestas aleatórias

O modelo de florestas aleatórias mostrou-se preciso na regressão e versátil, haja vista que foi possível realizar com ele testes que não são possíveis utilizando o modelo NRTL. Em quase todos os casos estudados, os coeficientes de determinação ajustados mostraram uma correlação forte entre os dados experimentais e calculados. As distribuições dos resíduos dos casos estudados com este modelo em todos os casos se aproximou de uma normal padrão, o que mostra que este é um modelo viável para estudar o equilíbrio líquido-líquido entre água, ácido acético e éster.

Além disso, foi possível realizar a predição dos dados de teste considerando os ésteres isômeros e todos os ésteres com bastante precisão. Observou-se uma queda de performance no estudo do modelo com todos os ésteres que não foi observada no estudo dos pares de isômeros, talvez porque aquele possuísse um conjunto de dados com propriedades muito

heterogêneas, o que pode ter atrapalhado a máquina no reconhecimento de um padrão matemático.

Uma melhor performance deste modelo poderia ser alcançada ao se utilizar um conjunto de dados maior, pois desta forma é possível utilizar uma porcentagem de dados para treinamento menor que a utilizada neste trabalho – que foi de 60% em todas as modelagens - e, mesmo assim, realizar um treinamento satisfatório do modelo.

6 CONCLUSÕES

O modelo NRTL já é consagrado na modelagem do equilíbrio líquido-líquido, mas o equacionamento, no qual este modelo se insere precisa ser revisto. Suposições de idealidade podem ter afetado os resultados. Uma alternativa ao equacionamento utilizado é o proposto por Prausnitz (1999), o qual pode ser aplicado a dados obtidos de forma experimental em trabalhos futuros.

O modelo de florestas aleatórias mostrou-se eficaz e mais preciso que o modelo NRTL na predição das frações molares dos componentes nas misturas estudadas por ambos os modelos. Além disso, há a possibilidade de se realizar modelagens de equilíbrio utilizando compostos semelhantes como se fossem um só sem impacto significativo na performance deste modelo, o que é particularmente interessante ao se considerar a aplicação deste modelo à misturas complexas, como o biodiesel, onde compostos semelhantes são majoritários na mistura. Esta vantagem não se observa ao se realizar a modelagem utilizando dados de compostos com propriedades muito distintas.

Por fim, conclui-se este trabalho com a constatação de que os modelos baseados em aprendizado de máquina possuem grande potencial de aplicação na termodinâmica do equilíbrio de fases. Estes modelos contornam alguns problemas daqueles baseados em energia de Gibbs em excesso - tais como Wilson, NRTL e UNIQUAC - que dependem de um bom equacionamento matemático para sua implementação e aumentam a complexidade quanto mais compostos estiverem presentes no sistema. Os algoritmos de regressão por aprendizado de máquina, além de serem mais diretos possuem implementação mais simples.

REFERÊNCIAS

- ATKINS, P. **Físico-química – Fundamentos**, Rio de Janeiro: LTC, 2001. 476 p.
- CARNITI, P.; CORI, L. RAGAINI, V. A Critical Analysis of the Hand and Othmer-Tobias Correlations. **Fluid Phase Equilibria**, v. 2, p. 39-47, 1978.
- CREMASCO, M. A. **Fundamentos de transferência de massa**, Campinas: Editora Unicamp, Campinas, 2002, 460 p.
- INCE, E.; KIRBASLAR, S. I. Liquid-liquid equilibria of the water-acetic acid-butyl acetate system. **Brazilian Journal of Chemical Engineering**, v. 19, n. 2, p. 243 - 254, 2002.
- MADEIRA, A. A. Application of a ternary phase diagram to the liquid-liquid extraction of ethanoic acid using ethyl ethanoate. **Eclética Química Journal**, v. 44, n. 3, p. 43-49, 2019
- MICHELSSEN, M.L.; MOLLERUP, J.M. **Thermodynamic Models: Fundamentals & Computational Aspects**. Lyngby: Tie-Line Publications, 2007. 382 p.
- OTHMER, D. F.; TOBIAS, P. E. Tie Line Correlation. **Industrial Engineering Chemistry**, v. 34, p. 693, 1942.
- PRAUSNITZ, J.M.; LICHTENTHALER, R. N; AZEVEDO, E. G. DE. **Molecular Thermodynamics of Fluid-Phase Equilibria**, New Jersey: Prentice-Hall, 1999. 860 p.
- PROCHAZKA, J. HEYBERGER, A. Correlation of ternary liquid-liquid equilibria in system isobutyl acetate-acetic acid-water. **Chemical Engineering Science**, v. 51, n. 6, p. 893-903, 1996
- RENON, H.; PRAUSNITZ, J. M. Local compositions in thermodynamic excess functions for liquid mixtures, **AIChE journal**, v. 14, n. 1, p. 135–144, 1968.
- ROQUE, L. R.; PINTO, R. R.; OLIVEIRA, L. H. Liquid-liquid equilibrium data for ternary systems of water + acetic acid + acetate esters at 293.2 K and 303.2 K and 95 kPa. **Fluid Phase Equilibria**, v. 463, p 34-47, 2018

TERRON, L. R. **Termodinâmica química aplicada**, Barueri: Manole, 2009. 876 p.

WANG, L., CHENG, Y., LI, X. Liquid-Liquid Equilibria for the Acetic Acid + Water + Amyl Acetate and Acetic Acid + Water + 2-Methyl Ethyl Acetate Ternary Systems, **Journal of Chemical and Engineering Data**, v. 52, n. 6, p. 2171 - 2173, 2007