Universidade Federal
do Rio de Janeiro
Escola Politécnica

ANÁLISE ESTATÍSTICA E DETERMINAÇÃO EXPERIMENTAL DE ÂNGULOS DE CONTATO E MOLHABILIDADES DE SISTEMAS ÓLEO-SALMOURA-MINERAL ANÁLOGOS AO PRÉ-SAL BRASILEIRO

Felippe Pardini de Souza

Projeto de Graduação submetido ao Corpo Docente do Curso de Engenharia de Petróleo da Escola Politécnica da Universidade Federal do Rio de Janeiro como parte integrante dos requisitos necessários à obtenção do título de Engenheiro de Petróleo.

Orientador: Prof. Santiago Gabriel Drexler, M. Sc.

Coorientador: Prof. Paulo Couto, Dr. Eng.

Rio de Janeiro

Março de 2018

ANÁLISE ESTATÍSTICA E DETERMINAÇÃO EXPERIMENTAL DE ÂNGULOS DE CONTATO E MOLHABILIDADES DE SISTEMAS ÓLEO-SALMOURA-MINERAL ANÁLOGOS AO PRÉ-SAL BRASILEIRO

Felippe Pardini de Souza

PROJETO DE GRADUAÇÃO SUBMETIDO AO CORPO DOCENTE DO CURSO DE ENGENHARIA DE PETRÓLEO DA ESCOLA POLITÉCNICA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE INTEGRANTE DOS REQUISITOS NECESSÁRIOS À OBTENÇÃO DO TÍTULO DE ENGENHEIRO DE PETRÓLEO.

Examinado por:

_____

Prof. Santiago Gabriel Drexler, M. Sc.

_____

Prof. Paulo Couto, Dr. Eng.

_____

Prof. Virgílio José Martins Ferreira Filho, D. Sc.

_____

William Godoy de Azevedo Lopes da Silva, M. Sc.

Rio de Janeiro

Março de 2018

Pardini de Souza, Felippe

Análise Estatística e Determinação Experimental de Ângulos de Contato e Molhabilidades de Sistemas Óleo-Salmoura-Mineral Análogos ao Pré-Sal Brasileiro / Felippe Pardini de Souza – Rio de Janeiro: UFRJ / Escola Politécnica, 2018.

XIX, 119p.:il.; 29,7cm.

Orientador: Santiago Gabriel Drexler

Coorientador: Paulo Couto

Projeto de Graduação – UFRJ / Escola Politécnica / Engenharia de Petróleo, 2018.

Referências Bibliográficas: p. 80-83.

1. Molhabilidade. 2. Ângulo de Contato. 3. Pré-Sal Brasileiro. 4. DSA. 5. Análise Estatística.

I. Drexler, Santiago Gabriel & Couto, Paulo. II. Universidade Federal do Rio de Janeiro, Escola Politécnica, Curso de Engenharia de Petróleo. III. Título.

Dedico este trabalho à minha família e a todos que de alguma forma contribuíram à minha formação profissional e pessoal durante estes cinco anos.

Agradecimentos

Gostaria de agradecer profundamente aos meus pais José Antonio Batista de Souza e Christiane Pardini de Souza, à minha irmã Christine Pardini de Souza e à minha melhor amiga, companheira e namorada Camilla Harumi Orlandi Ito. Sem o apoio e incentivo constantes seus, todo o período de minha graduação teria sido diferente e, sem dúvida, muito mais complicado. Palavras não descrevem o quanto eu amo vocês.

Agradeço ao meu orientador Prof. M. Sc. Santiago Drexler por todo o apoio e direcionamento ao longo da execução deste trabalho e também pelo conhecimento transmitido a mim e aos meus colegas de turma. Tenho grande apreço por tudo que aprendi com você. Sua didática, profissionalismo e vontade de ensinar são fascinantes e exemplares. Suas aulas marcaram-me profundamente como provas cabais de minha paixão pela Engenharia e pela Indústria Petrolífera. Admiro-o muito como profissional e pessoa.

Agradeço ao meu coorientador Prof. Dr. Paulo Couto por suas conquistas e intensas lutas como coordenador do curso de Engenharia de Petróleo na Poli UFRJ e também pelo vasto conhecimento transmitido em diversas disciplinas. Sua ajuda e seus direcionamentos ao longo desses cinco anos foram fundamentais em diversas situações.

Obrigado ao Prof. Dr. Virgílio pela oportunidade de Iniciação Científica no LORDE/COPPE/UFRJ ainda no segundo semestre da graduação, por ter aceito fazer parte da banca e por ter tido papel fundamental na construção do curso de Engenharia de Petróleo na UFRJ.

Obrigado M. Sc. William Godoy por ter aceito fazer parte de minha banca examinadora e principalmente por ser o epítome de uma pessoa inteligente e extremamente humilde. As breves mas proveitosas conversas que tive com você, que tinham assuntos variando de livros de Carl Sagan e Erwin Schrödinger até futilidades da vida, serão lembradas como provas de seu caráter curioso, esforçado e descontraído.

Agradeço aos meus companheiros e companheira do “PetroTeam UFRJ” (Ana Jerdy, Claudio Salinas, Elton Correia, Felipe Adrião, Felipe Pontual, Gabriel Magalhães e Vinícius Veleda), por dois anos fantásticos de muitas conquistas (em especial o histórico campeonato mundial do *PetroBowl*® *Championship* em 2016) e também de muito conhecimento e respeito mútuo. Aprendi com vocês a perfeita definição de equipe.

Lembrar-me-ei sempre das experiências que tive com vocês como algumas das melhores que já vivi e me sentirei sempre honrado e orgulhoso por ter sido capitão da equipe em 2017 e por ter representado a UFRJ e o Brasil ao redor do globo.

Obrigado também a Caio Dutra, Rafael Augusto B. R. Alves, à Ludmyla Carolina Barbosa e Mariana Frias de Albuquerque pela amizade e por todos os momentos que passamos juntos no Centro de Tecnologia e na Universidade, assim como nas Choppadas.

Obrigado ao Laboratório de Recuperação Avançada de Petróleo (LRAP) da COPPE/UFRJ e à Petrobras por proverem a infraestrutura e os recursos financeiros necessários para a realização dos experimentos e análises que compõem este trabalho. Obrigado também a Alex Gomes, Eng. Nathália Dias e M. Sc. Thaís Silveira pela amizade e pela valiosa ajuda na realização dos experimentos.

Por fim, gostaria de agradecer à UFRJ e Escola Politécnica pelos cinco anos mais produtivos, interessantes e intensos de minha vida, nos quais sinto que o que aprendi permitiu-me caminhar um passo além na curva de aprendizado da vida. E tenho a certeza de que continuarei caminhando incessantemente na busca pelo ponto crítico de primeira derivada nula e segunda derivada negativa desta curva.

“Para ser grande, sê inteiro: nada
Teu exagera ou exclui.
Sê todo em cada coisa. Põe quanto és
No mínimo que fazes.
Assim em cada lago a lua toda
Brilha, porque alta vive.”

Ricardo Reis, 1933

Resumo do Projeto de Graduação apresentado à Escola Politécnica/UFRJ como parte dos requisitos necessários para a obtenção do grau de Engenheiro de Petróleo.

ANÁLISE ESTATÍSTICA E DETERMINAÇÃO EXPERIMENTAL DE ÂNGULOS DE CONTATO E MOLHABILIDADES DE SISTEMAS ÓLEO-SALMOURA-MINERAL ANÁLOGOS AO PRÉ-SAL BRASILEIRO

Felippe Pardini de Souza

Março de 2018

Orientador: Prof. Santiago Gabriel Drexler, M. Sc.

Coorientador: Prof. Paulo Couto, Dr. Eng.

Curso: Engenharia de Petróleo

O ângulo de contato, fortemente relacionado à molhabilidade de sólidos, é um parâmetro de grande importância para a Engenharia de Reservatórios de hidrocarbonetos, visto que afeta diretamente o escoamento de fluidos em meios porosos e, consequentemente, o fator de recuperação de um campo petrolífero. A relevância aumenta ainda mais quando os campos são de caráter estratégico, como é o caso dos reservatórios do pré-sal brasileiro, os quais contêm impressionantes volumes de óleo de médio-alto grau API. O presente estudo foi conduzido para investigar a influência de três fatores (mineral, tempo de envelhecimento e solvente) nos resultados de ângulo de contato de sistemas constituídos por: óleo morto de um importante campo do pré-sal brasileiro; salmoura de alta salinidade (aprox. 226.000 ppm), que reproduzia precisamente em laboratório a salmoura de formação do mesmo reservatório em que o óleo foi retirado; e por um substrato (*i.e.* amostra de mineral). As medições foram feitas por um equipamento de tecnologia óptica, o tensiômetro de gota pendente (DSA, na sigla em inglês), em condições de temperatura e pressão fixas (60 ºC e 1000 psi). A análise das centenas de dados experimentais obtidos foi realizada com o auxílio de ferramentas estatísticas (*e.g.* matriz de correlação, ANOVA, diagrama de efeitos de Pareto e regressão linear múltipla), que realizaram a mineração de dados e embasaram as interpretações dos resultados, agregando valor e credibilidade ao trabalho empírico.

Palavras-chave: Molhabilidade, Ângulo de Contato, Pré-Sal Brasileiro, DSA, Análise Estatística

Abstract of the Undergraduate Project presented to POLI/UFRJ as partial fulfillment of the requirements for the degree of Petroleum Engineer.

STATISTICAL ANALYSIS AND EXPERIMENTAL DETERMINATION OF CONTACT ANGLES AND WETTABILITIES OF OIL-BRINE-MINERAL SYSTEMS ANALOGOUS TO THE BRAZILIAN PRE-SALT

Felippe Pardini de Souza

March/2018

Advisor: Prof. Santiago Gabriel Drexler, M. Sc.

Co-advisor: Prof. Paulo Couto, Dr. Eng.

Course: Petroleum Engineering

The contact angle, strongly related to the wettability of solids, is a parameter of paramount importance to Reservoir Engineering since it directly affects the flow of fluids in porous media and, consequently, the recovery factor of an oil field. The relevance is even greater if the oil fields have a strategic character, such as the ones belonging to the Brazilian Pre-Salt, which have astonishing volumes of medium-high API grade oils. This study was conducted to assess the influence of three factors (mineral, aging time and solvent) on the contact angle experimental results of systems composed by: dead oil from an important Pre-Salt field; high salinity brine (roughly 226,000 ppm), which precisely reproduced in laboratory the formation brine; and a rock slab (*i.e.* a sample of mineral). The measurements were made by an optical technology equipment, the Drop Shape Analyzer (DSA), in fixed pressure and temperature conditions (60 ºC and 1000 psi). The analysis of the experimental data obtained was made with the use of statistical tools (*e.g.* correlation matrix, ANOVA, Pareto chart of effects and multiple linear regression), which performed the data mining and supported the interpretation of the results, adding value and credibility to the empirical work.

Keywords: Wettability, Contact Angle, Brazilian Pre-Salt, DSA, Statistical Analysis

Sumário

Lista de Imagens xii

Lista de Tabelas xv

Lista de Símbolos xvii

1. Introdução 1

1.1. Estrutura 3

1.2. Motivação e Objetivos 1

2. Tópicos em Estatística e Probabilidade 5

2.1. Estatística Descritiva 8

2.1.1. Medidas de Posição 11

2.1.2. Medidas de Dispersão 13

2.1.3. Medidas de Assimetria 15

2.1.4. Medidas de Curtose 16

2.2. Teoria da Probabilidade 17

2.2.1. Funções e Distribuições de Probabilidade 20

2.2.2. Valor Esperado, Variância e Covariância de Variáveis Aleatórias 24

2.3. Estatística Indutiva 25

2.3.1. Estatísticas, Estimadores e Distribuições Amostrais 26

2.3.2. Estimação de Parâmetros 28

2.3.3. Testes de Hipóteses 29

2.3.4. ANOVA 33

2.4. Regressão Linear 39

2.4.1. Regressão Linear Simples 39

2.4.2. Regressão Linear Múltipla 43

2.4.3. ANOVA utilizando Regressão 44

2.5. Diagrama de Pareto 46

3. Tópicos em Fenômenos de Superfície 48

3.1. Tensão Interfacial 48

3.1. Molhabilidade 52
3.2. Ângulo de Contato 54
4. Medição de Ângulos de Contato: Metodologia e Procedimentos 57
4.1. O *Drop Shape Analyzer* 57
4.2. Metodologia Experimental 58
4.3. Preparação de Amostras 65
5. Resultados Experimentais e Análise Estatística 67
6. Considerações Finais 78
7. Referências Bibliográficas 80
Anexo I – Statsoft Statistica 7.0 84
Anexo II – Distribuições Tabeladas 89
Anexo III – Medições Experimentais 90

Lista de Imagens

Figura 1 - Consumo de energia primária por combustível (esquerda) e constituintes da energia primária em porcentagem (direita). É nítido o crescente aumento na demanda por energia primária e também o aumento da participação das energias renováveis e do gás natural. Fonte: Adaptado de [2]. 2
Figura 2 - Esquema para compreensão dos conceitos fundamentais da Estatística de acordo com o caso-exemplo apresentado. “R” representa “Reservatório”. 7
Figura 3 - Indução vs. Dedução. 8
Figura 4 - Exemplo de descrição gráfica de variáveis através de um gráfico de barras. 9
Figura 5 - Exemplo de descrição gráfica de variáveis através de um gráfico de setores. 10
Figura 6 - Histograma construído a partir dos intervalos da variável quantitativa e contínua VOIP. 11
Figura 7 - Exemplo de assimetria numa distribuição Gaussiana. Fonte: [6]. 16
Figura 8 - Comparação entre distribuições com diferentes curtoses. Fonte: [7]. 17
Figura 9 - A variável aleatória como função. *A* é um evento do espaço amostral S e *X* é a variável aleatória. 20
Figura 10 - Esquema representando a variável aleatória *X* e a função de probabilidade de *X*. 21
Figura 11 - Exemplo gráfico da definição de função densidade de probabilidade e do cálculo da probabilidade. Fonte: Adaptado de [12]. 23
Figura 12 - O processo de amostragem em uma população. 26
Figura 13 - Um testemunho rochoso. As pequenas fendas circulares são *plugs* que foram retirados como amostras. Fonte: [13]. 27
Figura 14 - Variação da CCO de acordo com os valores críticos. 31
Figura 15 - NPP gerado no *software Statistica 7.0*® a partir dos dados de demanda por campanhas sísmicas (população). Supõe-se que sigam uma distribuição Normal para aplicação da ANOVA. Através de uma análise rápida do gráfico, é possível observar que a distribuição de probabilidades é aproximadamente Normal. 38
Figura 16 - Representação gráfica do modelo de Regressão Linear Simples. Observe a reta de regressão populacional. 41
Figura 17 - Exemplo (fictício) de Diagrama de Pareto para o controle de qualidade em uma fábrica de dutos, relacionando defeitos diversos com a quantidade observada de cada um. 46
Figura 18 - Exemplo de um Diagrama de Efeitos de Pareto para um estudo de tensão interfacial (IFT) alcanos-água considerando três variáveis independentes: temperatura,

composição do alcano e pressão. Note que a temperatura é a variável mais influente no resultado de IFT, com significância estatística. 47
Figura 19 - Comparação entre gotas d'água muito pequenas (B) e muito grandes (D). Fonte: [25]. 48
Figura 20 – Esquemas do desbalanço de forças intermoleculares e criação de interface entre fluidos. Fontes: Adaptado de [28], [29] e [30]. 49
Figura 21 - Exemplo de uma molécula apolar, o gás cloro (Cl_2), e polar, o ácido clorídrico (HCl). Entre o átomo de hidrogênio e átomo de cloro há uma grande diferença de eletronegatividade, o que gera um momento de dipolo (μ_0) na molécula de HCl. Fonte: Adaptado de [31]. 51
Figura 22 - Esquema de um reservatório molhável à água. Fonte: [28]. 53
Figura 23 – Esquema ilustrando ângulos de contato (θ) e molhabilidades correspondentes. Considere o óleo imerso em um ambiente de salmoura. 55
Figura 24 - Ângulo de contato e equilíbrio de forças na interface entre os fluidos. A cor preta simboliza o óleo e a cor cinza, a salmoura. 55
Figura 25 - Métodos diretos (a e b) e indiretos (c e d) de medição experimental de ângulos de contato. Fonte: [30]. 56
Figura 26 - *Drop Shape Analyzer DSA100HP*® da Krüss (acima) e um exemplo de medição de gota sessil utilizando o *software DSA4*® e o Método da Tangente (abaixo). Fonte: Adaptado de [30] e [34]. 57
Figura 27 - Esquema dos componentes do DSA, acrescido de equipamentos auxiliares (Bombas Quizix® e recipientes de fluidos). 58
Figura 28- A imagem da esquerda ilustra a visão no *software DSA4*® da câmara de observação. É possível observar a linha de injeção no canto inferior da imagem. A imagem central ilustra a injeção de uma gota de óleo. A imagem da direira mostra um quartzo e o óleo em contato com o mineral, imersos em salmoura e prontos para a medição. 59
Figura 29 - Ajuste da *baseline* (linha azul) no *software* DSA4. O quadrado vermelho representa a área focal: no caso, buscava-se medir o ângulo direito da gota pendente de óleo imersa em no ambiente de salmoura. 59
Figura 30 - A imagem da esquerda representa o início de medição de ângulos de contato de um substrato de dolomita com óleo, imersos em salmoura. A imagem da direita ilustra o mesmo sistema ao final das medições. Observe a diferença de formato da gota e de ângulo de contato. O ângulo de contato médio no início da medição era de 59,4º, enquanto que ao final era de 63,7º. 60
Figura 31 - Esquema ilustrativo dos conceitos estatísticos envolvidos no trabalho. A partir de uma população de ângulos de contato (todos os ângulos de contatos

possíveis), selecionam-se subpopulações de ACs, ou seja, subconjuntos dos valores possíveis (subpopulações). Estes subconjuntos são os vinte e quatro experimentos (simbolizados por “E”) realizados, cada um com suas características (minerais, solventes e envelhecimento). Dentre cada subpopulação, foram selecionadas vinte medições (amostra). A figura ilustra o processo para o experimento 18 (E18). Vinte medições em cada um dos vinte e quatro experimentos totalizaram 480 medições. .. 63
Figura 32 - O processo de envelhecimento de amostras. 65
Figura 33 – Um *rock slab* cru, sem envelhecimento (esquerda). Uma amostra após o envelhecimento em óleo, limpeza com solvente e banho de salmoura (centro). O *DSA100HP*® e o início de um experimento (direita). 66
Figura 34 - Gota de óleo espalhada em uma amostra de quartzo limpa com n-hexano. 67
Figura 35 - Matriz de Correlação no formato de gráficos de dispersão. 70
Figura 36 - NPP da população de 460 ângulos de contato determinados experimentalmente. A linha reta indica ajuste perfeito à distribuição Normal. 71
Figura 37 - Verificação da normalidade dos resíduos da Regressão Linear Múltipla. A linha reta é o ajuste perfeito à distribuição Normal. Observa-se apresença de alguns *outliers* nos extremos. 73
Figura 38 - Reta ajustada pela Regressão Linear Múltipla e medições experimentais (pontos). 75
Figura 39 - Diagrama de Efeitos de Pareto, construído para α = 5%. 76
Figura 40 - Geração de rótulos e seus equivalentes numéricos para o *input* no *software*. 84
Figura 41 - *Inputs* para a *One-Way ANOVA* realizada na seção 2.3.4. 84
Figura 42 - Definição da variável dependente e independente. 85
Figura 43 - Resultado da ANOVA. Repare no nível de significância utilizado (em destaque): 5% 85
Figura 44 - Parte da tabela de *input*. 86
Figura 45 – Exemplo de exclusão de experimentos (subpopulações) cujas variâncias não respeitavam a regra prática da ANOVA. 86
Figura 46 – Iniciando a Regressão Linear Múltipla e ANOVA por Regressão Linear Múltipla. 87
Figura 47 - Seleção da Regressão Linear Múltipla. 87
Figura 48 - Resultado completo da Regressão Linear Múltipla. 88
Figura 49 - Valores tabelados da distribuição F de Snedecor considerando um nível de significância de 5%. V1 são os graus de liberdade no numerador e V2 os graus de liberdade no denominador. Fonte: [9]. 89

Lista de Tabelas

Tabela 1 - Exemplo de Volumes de Óleo In Place (VOIP) para os reservatórios do caso-exemplo. 10
Tabela 2 - Intervalos (classes de frequência) da variável VOIP, pontos médios e frequências absolutas. 11
Tabela 3 - Medidas de posição da variável VOIP. 12
Tabela 4 - Medidas de dispersão da variável VOIP. 14
Tabela 5 - Medidas de assimetria para os valores da variável VOIP do caso-exemplo. 16
Tabela 6 - Probabilidades do exemplo de lançamento de dado. 22
Tabela 7 - O Teste de Hipóteses e as possíveis configurações. 30
Tabela 8 - Função característica de operação com variação nos valores críticos. 31
Tabela 9 - Conjunto de dados históricos de demanda por campanhas sísmicas. 37
Tabela 10 - Tabela ANOVA para o exemplo. 38
Tabela 11 - Tabela ANOVA utilizando Regressão Linear Múltipla. 45
Tabela 12 – Características principais da salmoura de formação reproduzida em laboratório. Fonte: [36]. 61
Tabela 13 - Características principais do óleo morto do reservatório do pré-sal brasileiro. Fonte: [36]. 61
Tabela 14 - Experimentos de ângulo de contato realizados. O número do experimento não reflete ordem cronológica de realização: os experimentos foram realizados em ordem aleatória. 62
Tabela 15 - Substratos rochosos e seus respectivos pontos de zero carga. 63
Tabela 16 - Solventes e suas respectivas constantes dielétricas. 64
Tabela 17 - Variáveis termodinâmicas dos testes. 64
Tabela 18 - Medidas de posição, dispersão, assimetria e curtose da população de ângulos de contato. 67
Tabela 19 – Resultados experimentais de ângulo de contato. 68
Tabela 20 - Matriz de correlações dos dados experimentais. Valores em negrito indicam significância estatística considerando α = 5%. Nomenclatura: T_{es} é o tempo de envelhecimento em salmoura, T_{eo} o tempo de envelhecimento em óleo e *AC* é o ângulo de contato. 69
Tabela 21 - Variância das subpopulações de ângulos de contato. Os campos em verde simbolizam os experimentos que respeitam a regra prática de variâncias com base na variância do Experimento 2. 72
Tabela 22 - Tabela ANOVA via Regressão para os dados experimentais. α = 5%. 73

Tabela 23 - Dados experimentais obtidos (20 medidas estáveis para cada experimento). 90

Lista de Símbolos

ANOVA Análise de Variância;

A Amplitude;

A Energia Livre de Helmholtz;

A Evento Arbitrário “A”;

$\bar{A}$ Evento Contrário Arbitrário “A”;

a_3 Coeficiente de Assimetria;

$a_{Pearson}$ Coeficiente de Assimetria de Pearson;

a_4 Coeficiente de Curtose;

ANP Agência Nacional do Petróleo, Gás Natural e Biocombustíveis;

BI *Business Intelligence*;

CCP Coeficiente de Correlação de Pearson;

CCO Curva Característica de Operação;

C Valor Crítico;

DOF *Degrees of Freedom*;

DSA *Drop Shape Analyzer*;

EOR *Enhanced Oil Recovery;*

EMV Estimador de Máxima Verossimilhança;

e_o Limite do Intervalo de Confiança;

e_i Resíduo da Observação *i*;

E(X) Valor Esperado de X;

F Estatística de Teste F de Snedecor;

f Frequência;

f’ Frequência Relativa;

fp Função de Probabilidade;

f(x) Função Densidade de Probabilidade;

H_0 Hipótese Nula;

H_1 Hipótese Alternativa;

IOR *Improved Oil Recovery*;

IFT Tensão Interfacial;

$L(\mu_t)$ Função Característica de Operação do Teste de Hipóteses;

MANOVA Análise Multivariada de Variância;

MS *Mean Square*;

md Mediana;

MELNT Melhor Estimador Linear Não-Tendencioso;

m_i Momento de ordem *i*;

NNP *Normal Probability Plot*;

pzc Ponto de Zero Carga;

P(X) Probabilidade;

PDF *Probability Density Function* (Função Densidade de Probabilidade);

R^2 Coeficiente de Determinação;

$\mathbb{R}$ Conjunto dos Números Reais;

$\mathbb{R}_x$ Contradomínio da Variável Aleatória X;

S Espaço Amostral;

SQE Soma dos Quadrados (Entre Amostras);

SQR Soma dos Quadrados (Resíduos, Dentre Amostras);

SQT Soma dos Quadrados Totais;

SS *Sum of Squares*;

t Estatística de Teste de t de Student;

T_{es} Tempo de Envelhecimento em Salmoura;

T_{eo} Tempo de Envelhecimento em Óleo;

VPL Valor Presente Líquido;

VA Variável Aleatória;

V(X) Variância de X;

VOIP Volume de Óleo *In Place*;

$\hat{x}$ Estimador de *x* (Variável Genérica);

χ^2 Distribuição de Qui-Quadrado;

$\bar{x}$ Média Aritmética (Amostral);

z Estatística de Teste (Normal Reduzida).

Letras gregas

α Nível de Significância;

σ Desvio Padrão;

σ^2 Variância;

ε Constante Dielétrica;

γ Tensão Superficial/Interfacial;

μ Média Aritmética (Populacional);

μ_d Momento de Dipolo;

σ Desvio Padrão;

s^2 Variância (Amostral);

θ Ângulo de Contato;

ζ Perturbação Aleatória;

1. Introdução

1.1. Motivação e Objetivos

O estudo dos Fenômenos de Superfície, acoplado à análise de seus efeitos no escoamento de fluidos na escala de poro (*i.e.* escala micrométrica), são de suma importância para a Engenharia de Reservatórios de hidrocarbonetos e áreas correlatas. São esses fenômenos que determinarão, por exemplo, a relevância das pressões capilares no reservatório e, consequentemente, a distribuição de saturação de fluidos em seu domínio e também a maior facilidade ou dificuldade de se produzir os hidrocarbonetos comerciais contidos nele. Além disso, o estudo da complexa molhabilidade de sólidos (rochas, no caso da Indústria Petrolífera), que teve seu início em 1960 mas só recebeu a devida atenção no fim dos anos 80 [1], apresenta-se também como fundamental para o completo entendimento do escoamento e da recuperação de petróleo em reservatórios, visto que afeta, por exemplo, a permeabilidade relativa[1] dos fluidos. Até os dias atuais, a molhabilidade engloba um conjunto de fenômenos ainda não totalmente compreendido e se constitui como um campo de estudo em voga por causa dos avanços nos usos das técnicas de recuperação avançada (IOR, do inglês *Improved Oil Recovery)* e melhorada (EOR, do inglês *Enhanced Oil Recovery*) de petróleo.

Numa macro-escala, os fenômenos descritos anteriormente impactam diretamente em variáveis de muito interesse para os Engenheiros de Petróleo (*e.g.* fatores de recuperação e produção acumulada) e também para Companhias e Estados (*e.g.* receita e lucro). Otimizar a produção de hidrocarbonetos, aumentar o fator de recuperação dos campos e tornar a produção a mais sustentável a longo prazo possível são praticamente sinônimos de bons fluxos de caixa e maximização de Valores Presentes Líquidos (VPLs).

Além disso, otimizar a exploração de recursos não-renováveis garante um aproveitamento eficiente e eficaz das reservas de um país e também é uma das formas de suprir a crescente demanda por energia no mundo, a qual acompanha duas variáveis principais: o crescimento populacional e o desenvolvimento econômico-social dos

[1] Facilidade com que um meio se deixa atravessar por um fluido na presença de outros.

países. Com o constante crescimento populacional alavancado pelos países em desenvolvimento e com o aumento geral do poderio econômico mundial, a demanda por energia primária (*i.e.* aquela disponível na natureza, sem conversões ou transformações), em todos os cenários possíveis, apresenta previsões de aumento para o futuro iminente e longínquo (vide Figura 1).

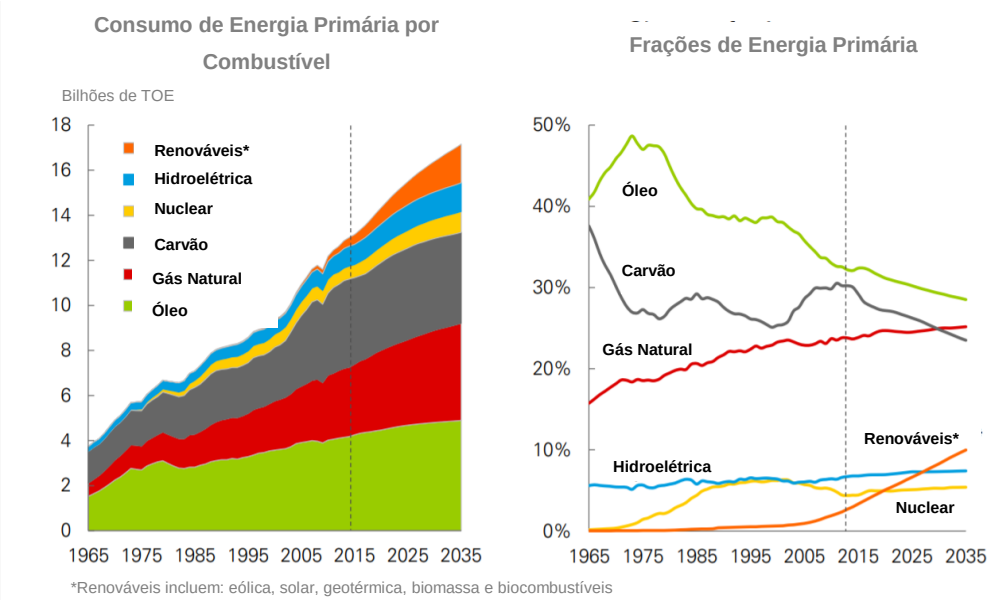

Figura 1 - Consumo de energia primária por combustível (esquerda) e constituintes da energia primária em porcentagem (direita). É nítido o crescente aumento na demanda por energia primária e também o aumento da participação das energias renováveis e do gás natural. Fonte: Adaptado de [2].

Com relação ao cenário da indústria petrolífera brasileira, o Pré-Sal é indiscutivelmente uma das descobertas mais importantes dos últimos tempos, tanto em volume quanto em qualidade, já que as reservas além de gigantescas contêm óleos considerados de boa qualidade (graus API médios ou altos). Portanto, realizar experimentos utilizando fluidos mortos (*i.e.* sem gás associado/dissolvido) reais (retirados de reservatórios do Pré-Sal) e minerais representativos do Pré-Sal, como é o caso deste trabalho, contribui para o melhor entendimento e exploração dos recursos que poderão tornar o Brasil a nova potência do mercado internacional de petróleo [3].

Quanto aos experimentos para determinação de ângulos de contato através do método da tangente, o qual é utilizado neste trabalho, não há normas que definam o processo de envelhecimento (o tempo que o sólido/mineral passa imerso em um fluido num recipiente sujeito à temperatura) e de limpeza (retirada do excesso de óleo da

superfície do sólido/mineral), ambos constituintes da fase pré-experimental de preparação de amostras. Logo, identificar as variáveis que mais afetam os resultados dos experimentos através da fundamentação matemática da Estatística, que se prova uma poderosa ferramenta de análise de dados e tomada de decisões na Engenharia de Reservatórios e na indústria como um todo, agrega valor e importância ao trabalho empírico.

Tendo em mente os pontos discutidos acima, este trabalho tem dois objetivos principais. O primeiro é obter medidas experimentais de ângulos de contato para sistemas óleo-salmoura-mineral do Pré-Sal e com isso caracterizar a molhabilidade dos minerais. O segundo, não menos importante, é realizar análises estatísticas dos dados para: melhor compreensão dos fenômenos envolvidos; confirmação de padrões esperados segundo a teoria; construção de análises de sensibilidade para identificação das principais variáveis que afetam o resultado das medições; e obtenção de uma correlação experimental que relacione ângulo de contato (variável dependente) com as variáveis independentes (mineral, tempo de envelhecimento e solvente) dos experimentos.

1.2. Estrutura

O presente trabalho foi estruturado de forma a tornar a abordagem aos assuntos a mais suave e clara possível. Para tanto, foi dividido em seis seções (excluindo a seção destinada às Referências Bibliográficas) e de forma que a fundamentação teórica dos fenômenos e ferramentas envolvidos(as) no assunto deste projeto fosse apresentada antes dos resultados experimentais e da análise estatística.

A primeira seção pretende fornecer as motivações e objetivos gerais, contextualizando o campo de estudo tratado neste trabalho nas atuais conjunturas dos mercados nacional e internacional de petróleo, além de descrever a importância das análises para o meio acadêmico e industrial.

A segunda seção introduz e revisa conceitos e equações de Estatística e Probabilidade, os quais serão utilizados posteriormente nas análises dos dados experimentais obtidos (seção 5). Definições de Estatística Descritiva, Estatística Indutiva e Teoria da Probabilidade são fornecidas de modo a melhorar a compreensão das técnicas utilizadas na seção supracitada. O foco da seção 2 é expor a teoria básica e alguns exemplos práticos, não fornecer demonstrações matemáticas complexas. O(a)

leitor(a) que já é familiarizado(a) com os conceitos sentir-se-á confortável em iniciar a leitura na terceira seção.

A terceira seção define os princípios e equações matemáticas que regem os Fenômenos de Superfície, com foco nos conceitos de tensão superficial/interfacial, ângulos de contato e molhabilidade de sólidos.

A quarta seção detalha a metodologia e os procedimentos experimentais para a medição dos ângulos de contato e também os princípios do tensiômetro de gota pendente/séssil (ou DSA, do inglês *Drop Shape Analyzer*), equipamento utilizado em todas as medições.

A quinta seção compila os resultados obtidos pelo DSA e demonstra toda a análise estatística feita com os dados, interpretando-os com detalhe.

A sexta e última seção reúne os principais pontos abordados em todo o projeto para chegar a uma conclusão, que se utiliza dos *outputs* da análise estatística para determinar qual fator experimental (variável independente) é mais importante para os resultados de ângulo de contato, além de descrever possíveis trabalhos futuros.

2. Tópicos em Estatística e Probabilidade

A Estatística é a ciência que se preocupa com a organização, descrição, análise e interpretação dos dados experimentais, visando a tomada de decisões [4].

Com aplicações em qualquer área do conhecimento que trabalhe com dados e conjuntos de informações, a Estatística prova-se como ferramenta muito útil em mineração de dados (*i.e.* organização de dados de forma a encontrar padrões, associações ou anomalias), controle de qualidade em processos industriais, redes neurais, análise e gerenciamento de risco, logística e cadeia de suprimentos, *business intelligence* (BI), pesquisas quantitativas (*e.g.* censos), entre diversas outras aplicações.

Até o final século XIX, a quantidade de dados e medições não justificava o desenvolvimento de uma ciência que lidasse com sua modelagem e interpretação. Porém, principalmente após o início do século XX, tal ciência tornou-se indispensável e revolucionou o meio científico.

Há um experimento mental em forma de história, proposto por Sir Ronald Aylmer Fisher, eminente Estatístico e um dos nomes mais importantes da Estatística, que descreve os primórdios da ciência de forma interessante [5]:

"(...) em Cambridge, Inglaterra, no final dos anos 1920, um grupo de professores universitários, suas esposas e alguns convidados tomavam o chá da tarde em uma mesa de um jardim. Uma das mulheres insistia em afirmar que o chá servido sobre o leite parecia ficar com gosto diferente do que apresentava ao receber leite sobre ele. Um dos homens na mesa animou-se e quis testar a proposição através de um experimento, no qual a senhora que insistira haver diferença seria servida com uma sequência de xícaras, algumas com leite servido sobre o chá, e outras com o chá servido sobre o leite."

A história, à primeira vista, parece inocente. Não é trivial perceber que ela ilustra uma mudança de pensamento revolucionária: buscava-se testar hipóteses e comprová-las através de ferramentas matemáticas. O resultado de um experimento era importante, mas a partir daquela época a confiabilidade nos dados e o estudo dos erros e incertezas inerentes a qualquer experimento tornaram-se tão relevantes quanto o resultado.

Antes de qualquer outra definição ou equacionamento, faz-se necessário definir os quatro conceitos fundamentais da Estatística:

- Universo: conjunto de objetos com pelo menos uma característica de interesse;

- Variável: qualquer característica, numérica ou não, dos elementos do universo. Podem ser quantitativas (contínuas ou discretas) ou qualitativas;
- População: conjunto de valores de uma variável. O tamanho da população é igual ao tamanho do Universo;
- Amostra: subconjunto finito de uma população.

Para exemplificar esses conceitos, imagine o seguinte caso-exemplo na indústria do petróleo, o qual será utilizado nesta e nas próximas seções deste trabalho: uma operadora, *i.e.* a companhia que detém os direitos de exploração de um campo, deseja analisar e avaliar um campo recém-adquirido num leilão da ANP (Agência Nacional do Petróleo, Gás Natural e Biocombustíveis). O campo é composto por dez reservatórios, que possuem diferentes volumes de óleo *in place* (VOIP), diferentes áreas e podem ser classificados como “bons”, “médios” ou “ruins” em termos de projeções de produtividade. A empresa também deseja realizar estudos mais complexos no campo e, para facilitar a análise e diminuir o tempo/custo de simulações computacionais, selecionará aleatoriamente um reservatório de cada uma das classificações, totalizando três reservatórios (um bom, um médio e um ruim).

No caso descrito acima, o universo são todos os dez reservatórios do campo. As variáveis são: volume de óleo e área (ambas quantitativas e contínuas); classificação (qualitativa). Os conjuntos de valores dessas variáveis são as populações (dez volumes de óleo, dez áreas e dez classificações). Note que, conforme já destacado anteriormente, o tamanho das populações é igual ao tamanho do universo. As amostras são os três reservatórios escolhidos e os respectivos valores de suas variáveis. O exemplo é esquematizado na Figura 2, que leva em consideração somente a variável qualitativa “classificação”.

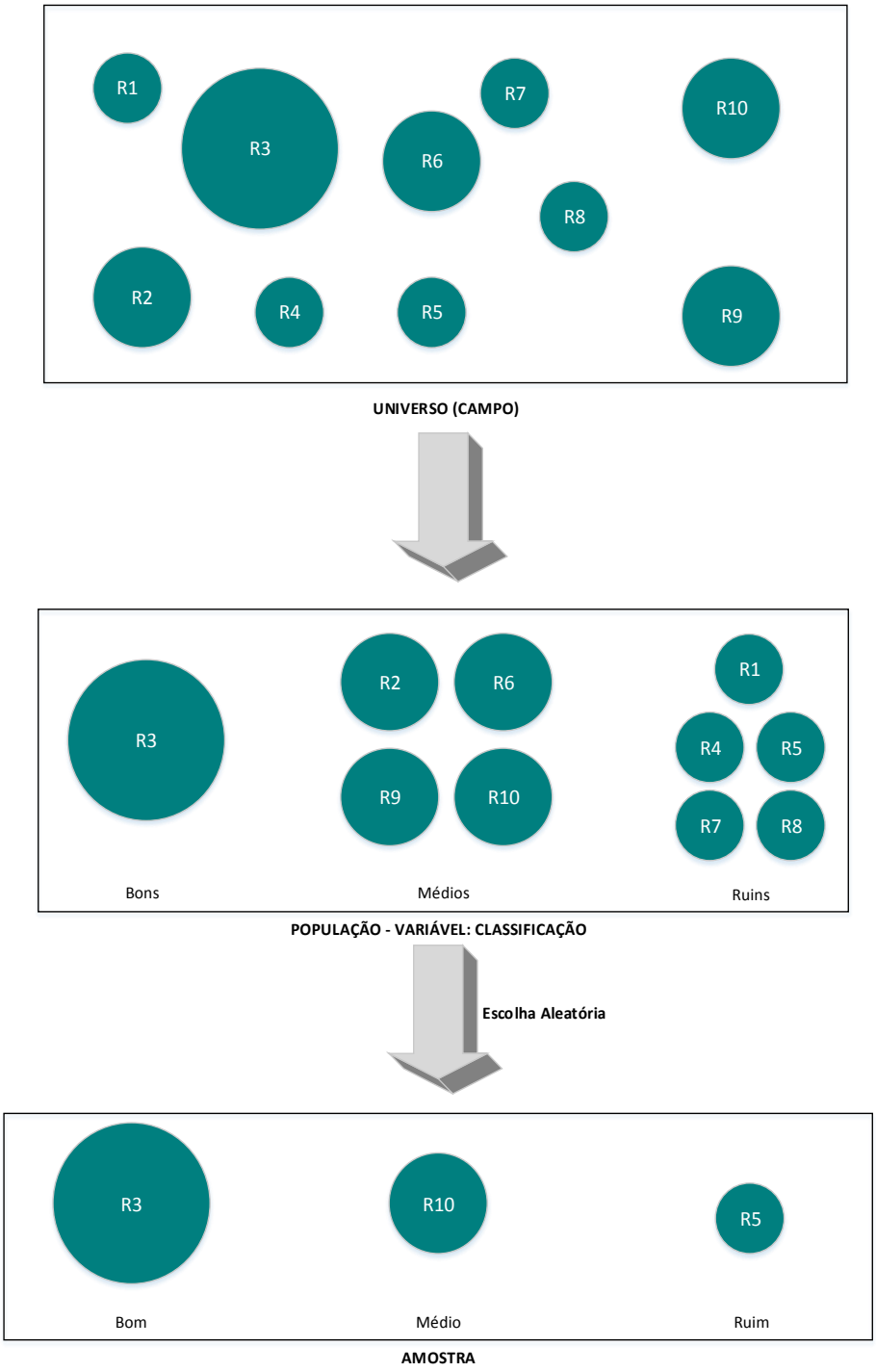

Figura 2 - Esquema para compreensão dos conceitos fundamentais da Estatística de acordo com o caso-exemplo apresentado. “R” representa “Reservatório”.

É importante destacar que os processos de definição da população e da amostra são, na maioria dos casos, complexos. Falhas durante esses processos geram conclusões erradas ou que não refletem com acurácia a realidade do fenômeno estudado.

A Estatística pode ser subdividida em duas vertentes: Estatística Descritiva e Estatística Indutiva. Enquanto a primeira organiza e descreve os dados de forma gráfica (*e.g.* histogramas) e/ou de forma numérica (*e.g.* medidas de tendência central, medidas de dispersão, momentos, curtose e assimetria), a última pretende tirar conclusões com base nos resultados observados em amostras extraídas de uma população e, para isso, envolve Probabilidade. É intuitivo pensar que quanto maior a amostra, mais confiável e precisa será a indução. O termo "indutiva" decorre da existência de um processo de indução, isto é, um processo de raciocínio em que, partindo-se do conhecimento de uma parte, procura-se tirar conclusões sobre o todo [4]. O termo é oposto ao processo de dedução (vide Figura 3).

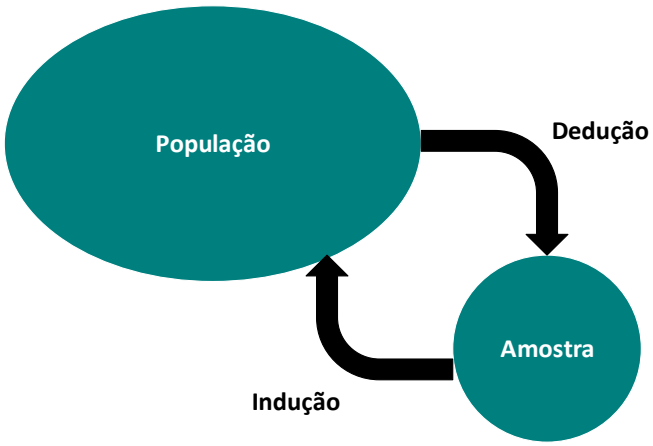

Figura 3 - Indução vs. Dedução.

2.1. Estatística Descritiva

As variáveis de um problema podem ser divididas em qualitativas, quando resultam de uma classificação por tipos ou atributos, e quantitativas, quando seus valores são expressos em números. As últimas podem ainda ser subdivididas em discretas, quando assumem somente valores pertencentes a um conjunto enumerável, e contínuas, quando podem assumir qualquer valor dentre um intervalo [4]. Exemplos destas definições foram dados no caso-exemplo.

A forma mais básica de se descrever um conjunto de variáveis é através da Descrição Gráfica. A frequência ou frequência absoluta (número de observações) de um dado valor/classificação de uma variável pode ser disposta em um dos eixos, enquanto

o outro pode conter o valor/classificação correspondente àquela frequência. Definindo a frequência *f* em forma de equação:

$$\boldsymbol{f} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}} \boldsymbol{f_i} = \boldsymbol{n} \ , \qquad (\mathbf{1})$$

onde f_i é a frequência do *i*-ésimo valor/classificação, *k* é o número de diferentes valores/classificações de uma variável e *n* o número total de elementos observados. Podemos definir também a frequência relativa, que é dada pelo quociente entre a frequência absoluta e *n*:

$$f'_i = \frac{f_i}{n} \ . \qquad (2)$$

Então, utilizando o caso-exemplo, podemos descrever os reservatórios e suas classificações através de gráficos de barras (Figura 4) e gráficos de setores (Figura 5).

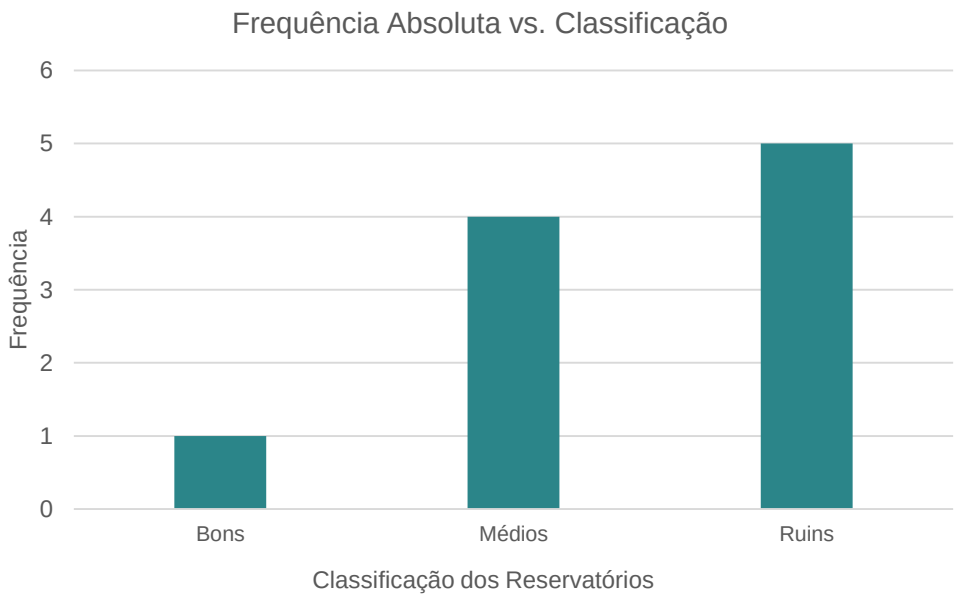

Figura 4 - Exemplo de descrição gráfica de variáveis através de um gráfico de barras.

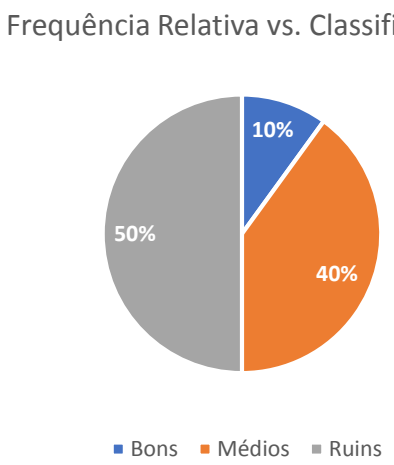

Figura 5 - Exemplo de descrição gráfica de variáveis através de um gráfico de setores.

Caso a variável seja qualitativa e contínua, a forma mais comum de representação gráfica é o Histograma, um gráfico de barras verticais cujas alturas correspondem às frequências de intervalos de valores da variável com amplitudes pré-definidas (Figura 6). Tomando a variável VOIP do caso-exemplo, tendo-se os dados presentes da Tabela 1, pode-se obter os pontos médios e as frequências absolutas como mostrado na Tabela 2.

Tabela 1 - Exemplo de Volumes de Óleo In Place (VOIP) para os reservatórios do caso-exemplo.

Reservatório	VOIP (em MMbbl*)
R1	5,2
R2	20,8
R3	35,1
R4	10,5
R5	10,6
R6	20,5
R7	10,8
R8	10,7
R9	20,0
R10	20,2

*Milhões de barris (1 bbl = 159 litros)

Tabela 2 - Intervalos (classes de frequência) da variável VOIP, pontos médios e frequências absolutas.

Classes de Frequência	Pontos Médios	Frequência Absoluta
[5,2-19,2]	12,2	5
(19,2-33,2]	26,2	4
(33,2-47,2]	40,2	1

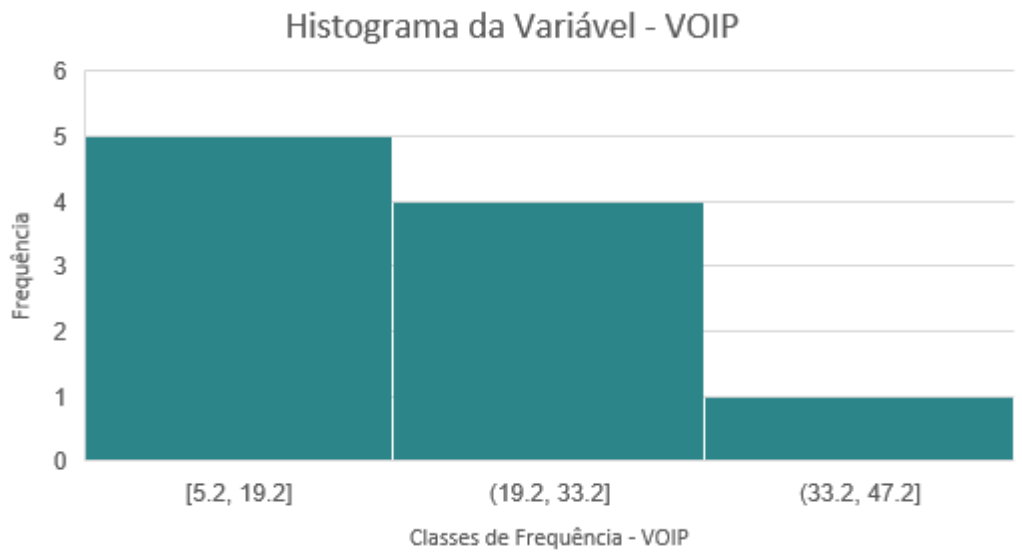

Figura 6 - Histograma construído a partir dos intervalos da variável quantitativa e contínua VOIP.

A outra forma de descrição de variáveis é a Descrição Numérica, possível somente para variáveis quantitativas. Para tanto, utilizam-se algumas medidas de posição, dispersão, assimetria e curtose (achatamento).

2.1.1. Medidas de Posição

O título destas medidas é autoexplicativo: localizar os valores de uma variável dentro do espectro de seus valores possíveis ou observados. São três as mais comuns: média aritmética, que é sensível a valores extremos e é dada por

$$\bar{x} = \sum_{i=1}^{k} f'_i x_i \ , \qquad (3)$$

onde $\bar{x}$ é a média aritimética de uma variável *x* qualquer e x_i é o i-ésimo valor dessa variável; a mediana, que não é sensível a valores extremos e pretende dividir o conjunto ordenado de valores em duas partes com igual número de elementos. Pode ser entendida também como o valor da variável em que a frequência acumulada (soma das frequências absolutas ou relativas) é igual a 50% e é dada, no caso de valores agrupados em classes, por

$$md = L_i + h_{md}\frac{\frac{n}{2} - F_a}{f_{md}}, \tag{4}$$

onde *md* é a mediana, h_{md} é a amplitude da classe que contém a mediana, *n* é o número de elementos do conjunto de dados, F_a é a soma das frequências das classes anteriores à que contém a mediana e f_{md} é a frequência da classe que contém a mediana [4]; e a moda ou classe modal, ou seja, valor ou a classe de maior frequência. Pode haver mais de uma moda em um conjunto de dados, configurando-o como bi, tri, quadrimodal e assim por diante.

Seguindo o caso-exemplo e a análise da variável VOIP, teríamos medidas de posição descritas na Tabela 3.

Tabela 3 - Medidas de posição da variável VOIP.

Média [MMbbl]	16,4
Mediana [MMbbl]	19,2
Classe Modal	[5,2-19,2]

É possível também generalizar a ideia da mediana para dividir o conjunto ordenado de valores em quatro subconjuntos com igual número de elementos, dando origem aos quartis (sendo *md* o 2º quartil), ou em cem subconjuntos, dando origem aos percentis. Se um valor de uma variável é o 80º percentil, por exemplo, isso significa que 80% dos outros valores desta mesma variável são menores que ele (ou, similarmente, que 20% dos outros valores são maiores que ele). Essas separatrizes são bastante utilizadas na literatura científica.

2.1.2. Medidas de Dispersão

As medidas de dispersão indicam o quanto os dados se apresentam afastados (dispersos) em relação à região central. Segundo COSTA NETO [4], são quatro as mais utilizadas: a amplitude, que é a diferença entre o maior e o menor valor do conjunto de dados, dada por

$$A = x_{max} - x_{min}, \tag{5}$$

onde A é a amplitude e x_{max}, x_{min} são o valor máximo e mínimo da variável, respectivamente; a variância (σ_x^2), dada por

$$\sigma_x{}^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n} \tag{6a}$$

ou, no caso de dados dispostos em uma tabela de frequências,

$$\sigma_x{}^2 = \sum_{i=1}^{n} \frac{f_i(x_i - \bar{x})^2}{n'}\,, \tag{6b}$$

onde σ_x^2 é a variância, n o número total de dados e $n' = n$, se n representar toda uma população ou $n' = n - 1$, se n representar uma amostra; o desvio-padrão, que é a raiz quadrada positiva da variância, ou seja,

$$\sigma_x = +\sqrt{\sigma_x{}^2} = +\sqrt{\sum_{i=1}^{n} \frac{f_i(x_i - \bar{x})^2}{n'}}\,, \tag{7}$$

onde σ_x é o desvio-padrão; e o coeficiente de variação, que é adimensional e definido como o quociente entre o desvio-padrão e a média, sendo geralmente expresso em porcentagem e dado por

$$C_v = \frac{\sigma_x}{\bar{x}} \ , \tag{8}$$

onde C_v é o coeficiente de variação.

É importante salientar que, no caso de dados agrupados em classes de frequência, como é comum em estudos que envolvem variáveis quantitativas e contínuas, o uso dos pontos médios das classes para o cálculo da variância tende a superestimá-la. Nos casos em que as classes têm mesma amplitude, o que é comum, a Correção de Sheppard (Eq. 9) é utilizada:

$$\sigma_{x,shep}{}^2 = \sigma_x{}^2 - \frac{h^2}{12} \ , \tag{9}$$

onde $\sigma_{x,shep}{}^2$ é a variância corrigida e h é a amplitude das classes.

No caso-exemplo, teríamos as medidas de dispersão apresentadas na Tabela 4. É interessante notar as unidades de cada uma.

Tabela 4 - Medidas de dispersão da variável VOIP.

Amplitude [MMbbl]	29,90
Variância [MMbbl²]	74,55
Desvio-Padrão [MMbbl]	8,63
Coef. De Variação	52,52%

A variância tem a unidade da variável estudada ao quadrado e tem pouca utilidade prática. O desvio-padrão sana esse inconveniente e o coeficiente de variação apresenta-se como boa medida para análises rápidas e preliminares sobre dispersão. É fácil, por exemplo, ver que o conjunto de valores da variável VOIP do caso-exemplo é bastante disperso quando se observa o C_v calculado na Tabela 4.

2.1.3. Medidas de Assimetria

Essas medidas têm o objetivo de caracterizar como e quanto o conjunto de dados e sua distribuição de frequência se afastam da condição de simetria [4]. São duas principais: o coeficiente de assimetria, adimensional e dado por

$$a_3 = \frac{m_3}{{\sigma_x}^3}\,, \tag{10}$$

onde a_3 é o coeficiente de assimetria e m_3 é o momento[2] centrado de terceira ordem em relação à média, ou seja,

$$m_3 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^3}{n}\,; \tag{11}$$

e o índice de assimetria de Pearson, também adimensional e definido como

$$a_{Pearson} = \frac{\bar{x} - m_0}{\sigma_x}, \tag{12}$$

onde $a_{Pearson}$ é o índice de assimetria de Pearson e m_0 é o momento centrado de ordem zero em relação à média, ou seja,

$$m_0 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^0}{n} = \sum_{i=1}^{n} \frac{1}{n}. \tag{13}$$

Um coeficiente de assimetria nulo significa que a distribuição de frequências é simétrica, enquanto que se for positivo indica assimetria positiva e se for negativo indica

[2] Os momentos da Estatística são importantes para caracterizar distribuições de probabilidade. Momentos de primeira, segunda, terceira e quarta ordem caracterizam, respectivamente, tendência central (*e.g.* média e mediana), dispersão (*e.g.* variância), assimetria e curtose (seção 2.1.4).

assimetria negativa. Um exemplo gráfico utilizando uma distribuição Gaussiana (Normal), muito utilizada na Probabilidade, é dado na Figura 7.

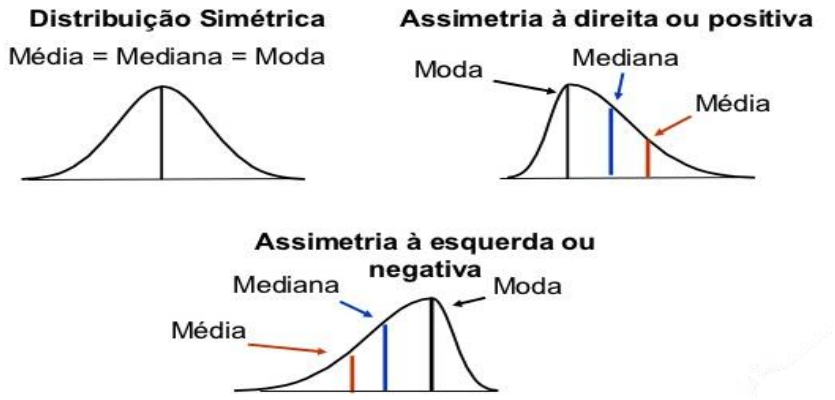

Figura 7 - Exemplo de assimetria numa distribuição Gaussiana. Fonte: [6].

Já com relação ao índice de assimetria de Pearson, se $|a_{Pearson}| < 0.15$, considera-se que a distribuição é praticamente simétrica. Se $0.15 < |a_{Pearson}| < 1$, a assimetria é considerada leve e se $|a_{Pearson}| > 1$, é considerada forte.

No caso-exemplo, temos as medidas de assimetria da Tabela 5. Nota-se que a distribuição dos valores da variável VOIP é assimétrica positiva segundo a_3 e fortemente assimétrica segundo $a_{Pearson}$.

Tabela 5 - Medidas de assimetria para os valores da variável VOIP do caso-exemplo.

Coef. de Assimetria	0,71
$a_{Pearson}$	1,78

2.1.4. Medidas de Curtose

As medidas de Curtose, ou Achatamento, caracterizam a forma da distribuição quanto a seu achatamento. Para classificação e comparação, utiliza-se a mesma distribuição Gaussiana citada na subseção anterior como base. Se a distribuição estudada é Gaussiana, ela é considerada mesocúrtica. Se ela for mais achatada que

uma Gaussiana, é considerada platicúrtica. E, finalmente, se for menos achatada que uma Gaussiana, é considerada leptocúrtica. A Figura 8 ilustra esses conceitos.

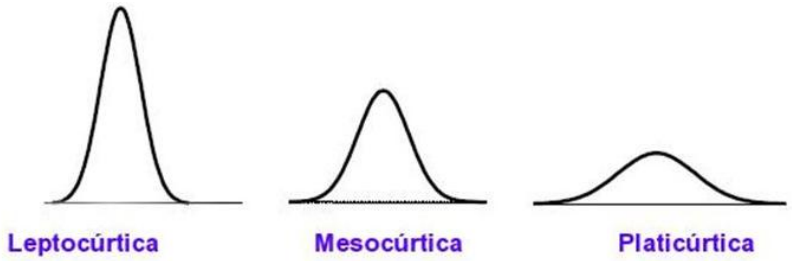

Figura 8 - Comparação entre distribuições com diferentes curtoses. Fonte: [7].

Segundo COSTA NETO [4], a caracterização do achatamento só tem sentido prático se a distribuição estudada for pelo menos aproximadamente simétrica. Das medidas de curtose, o coeficiente de curtose é o mais aplicado. O mesmo é adimensional e definido como o quociente do momento centrado de quarta ordem pelo quadrado da variância, subtraído de três:

$$a_4 = \frac{m_4}{{\sigma_x}^4} - 3 \tag{14}$$

onde a_4 é o coeficiente de curtose e m_4 é análogo à Equação 11, porém com a potência do numerador à quarta.

Um coeficiente menor que zero indica uma distribuição platicúrtica, igual a zero uma distribuição mesocúrtica, e maior que zero uma distribuição leptocúrtica.

O conjunto de dados de VOIP do caso-exemplo é altamente assimétrico e, portanto, não é relevante calcular uma medida de curtose para o mesmo.

2.2. Teoria da Probabilidade

A Estatística é construída com base em fundamentos da Probabilidade. São campos intimamente relacionados e complementares. A Estatística Indutiva (seção 2.3), em especial, utiliza-se de diversos conceitos da Teoria da Probabilidade, que serão apresentados nesta seção.

A Probabilidade tem uma longa e rica história que tem seu início em meados do século XVII, quando Chevalier de Meré, Pascal e Fermat desenvolveram formulações matemáticas para as chances em jogos de azar [8]. O grande desenvolvimento da teoria, porém, ocorreu no século XX e partiu do entendimento dos cientistas de que a grande maioria dos fenômenos observados na natureza não era determinística, ou seja, que dado um conjunto de condições um evento sempre ocorria, mas sim probabilísticos (estocásticos), ou seja, que dado um conjunto de condições um evento tem certa chance (probabilidade) de ocorrer. Desde então, a teoria evoluiu a ponto de conseguir modelar inúmeros fenômenos estocásticos de alta complexidade.

Assim como no caso da Estatística, faz-se necessário definir os conceitos fundamentais da probabilidade [4]:

- Aleatoriedade: relativo a um processo repetitivo cujo resultado não descreve um padrão determinístico;
- Experimento: conjunto de procedimentos que visam testar uma hipótese;
- Experimento aleatório: aquele que pode realizar-se repetida e indefinidamente e em que após grande número de repetições é possível observar regularidade na ocorrência dos eventos aleatórios;
- Espaço amostral: conjunto de resultados possíveis do experimento aleatório;
- Evento: acontecimento de interesse. Constitui-se como qualquer subconjunto do espaço amostral;
- Eventos independentes: aqueles em que a ocorrência de um não influencia na ocorrência do outro.

Há fundamentalmente duas definições formais de probabilidade. A definição clássica descreve que se um evento A puder ser decomposto em *m(A)* eventos favoráveis pertencentes a uma partição do espaço amostral (*S*) formada por *l* eventos igualmente prováveis, então a probabilidade do evento *A* [*P(A)*] é dada pela Equação 15 [9]:

$$P(A) = \frac{m(A)}{l}. \tag{15}$$

A outra definição, intitulada frequentista, descreve que a probabilidade é o limite (não-matemático) da frequência relativa de um evento qualquer (A) quando o número

de realizações (*n*) de um experimento tende ao infinito [10]. É nítida a relação da definição frequentista com a Estatística.

Como exemplo ilustrativo da definição clássica, tomemos uma plataforma de produção da Bacia de Campos do litoral brasileiro e quatro Engenheiros(as) de Petróleo embarcados na mesma. Consideremos que o espaço amostral são os dias de embarque dos Engenheiros(as). O(a) engenheiro(a) "1" está embarcado há 15 dias, o(a) "2" e "3" há 30 dias e "4" há 40 dias. Gostaríamos de saber qual a probabilidade de, escolhendo um dos quatro Engenheiros ao acaso (aleatoriamente), ele(a) estar embarcado(a) há 30 dias (evento A). Portanto, temos o espaço amostral de dias de embarque dado pela Equação 17.

$$S = \{15, 30, 30, 40\} \tag{17}$$

A probabilidade do evento A é de 50%, já que dos quatro Engenheiros(as) embarcados (*l*), dois estão na plataforma há 30 dias [*m(A)*] (vide Eq. 18).

$$P(A) = \frac{m(A)}{l} = \frac{2}{4} = 50\% \tag{18}$$

Através do exemplo, é intuitivo observar também que a probabilidade de escolhermos um engenheiro(a) que não está embarcado há 30 dias, ou seja, o contrário do evento A (simbolizado frequentemente por $\bar{A}$), é também de 50%, ou seja,

$$P(\bar{A}) = 1 - P(A) = 50\% \tag{19}$$

A Equação 19 descreve uma das diversas propriedades da Probabilidade para o caso de eventos aleatórios independentes. As demais propriedades não serão tratadas por não se enquadrarem no objetivo da presente seção e do trabalho em geral.

2.2.1. Funções e Distribuições de Probabilidade

As funções e distribuições de probabilidade aplicam-se às variáveis aleatórias (VA). Uma variável aleatória *X* é uma função cujo domínio é um subconjunto do espaço amostral (*i.e.* evento) ou ele todo e o contradomínio (imagem) é um subconjunto de $\mathbb{R}$ ($\mathbb{R}_x$) ou ele todo [8]. A Figura 9 ilustra o conceito.

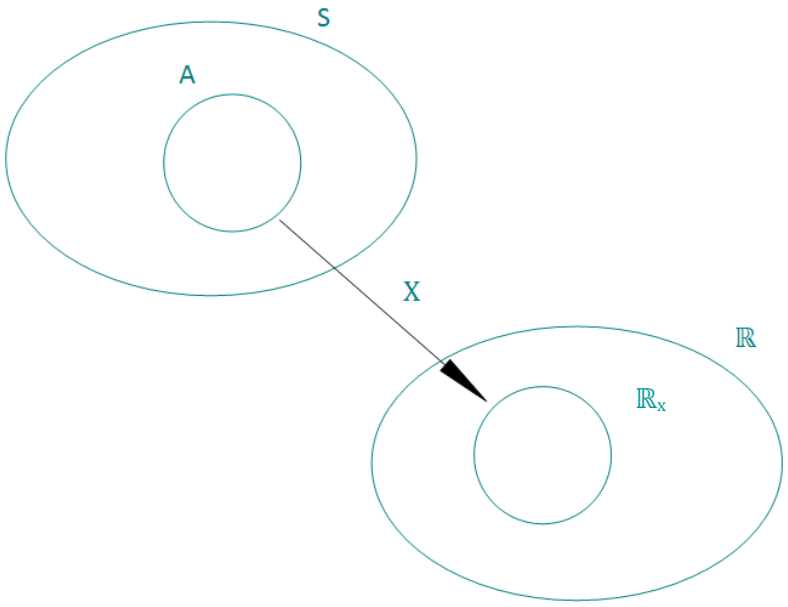

Figura 9 - A variável aleatória como função. *A* é um evento do espaço amostral S e *X* é a variável aleatória.

Se *X* é uma variável aleatória discreta e x_1, x_2, ... , x_i são seus valores possíveis, podemos associar a cada valor x_i sua probabilidade, conforme demonstra a Equação 20:

$$P(x_i) = P(\{s_i \in S \mid X(s) = x_i\}) , \tag{20}$$

onde s_i pertence ao domínio *S* e $P(x_i)$ é a função de probabilidade (fp) de *X.*

Os valores de $P(x_i)$ satisfazem às condições [9]:

$$P(x_i) \geq 0 , \tag{21}$$

e

$$\sum_{x_i \in \mathbb{R}_x} P(x_i) = 1\,. \tag{22}$$

O conjunto dos pares $\{x_i, P(x_i)\}$ é a distribuição de probabilidade de X. Graficamente, a Figura 10 esquematiza os conceitos acima.

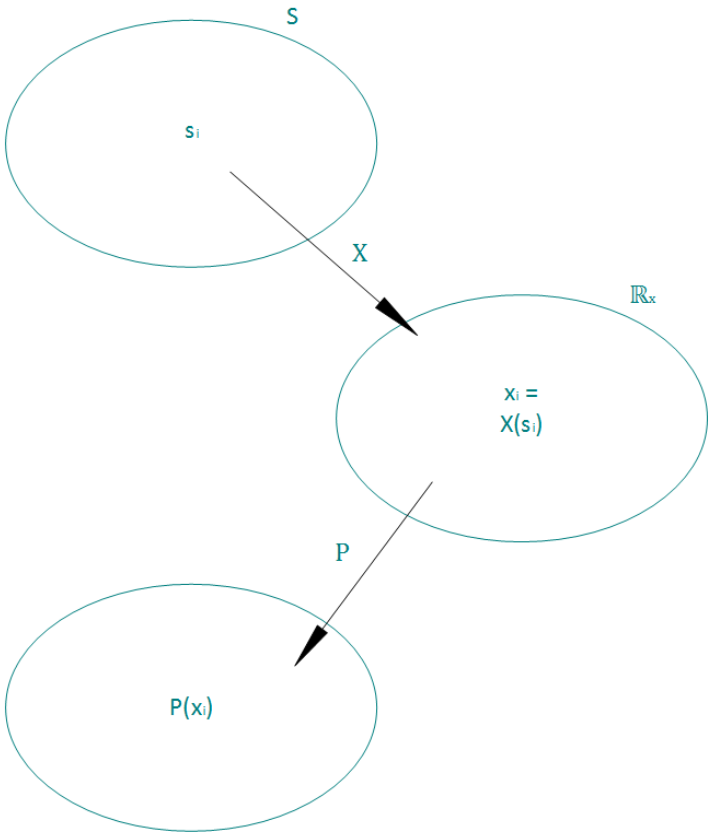

Figura 10 - Esquema representando a variável aleatória *X* e a função de probabilidade de *X*.

A distribuição de Binomial, por exemplo, é aplicável a VAs discretas e modela situações em que o experimento são *n* tentativas independentes de uma mesma ação e o evento de interesse é a quantidade “sucessos”. O clássico experimento do lançamento de um dado e a quantidade de resultados iguais a seis (variável aleatória *X*) são facilmente modelados pela distribuição Binomial [9]. Dizemos então que *X* tem distribuição Binomial com a seguinte notação:

$$X \sim Bi\ (n, p),$$

onde *Bi (n,p)* indica as seguintes informações: distribuição Binomial, número de tentativas *n*, com probabilidade de ocorrência do evento (“sucesso”) *p*.

A função de probabilidade da distribuição Binomial é dada pela Equação 23 [9].

$$P(X = k) = C_n^k p^k (1-p)^{n-k} \tag{23}$$

onde $k \in \mathbb{R}_x$ tal que $k = 0, 1, \ldots, n$.

No exemplo do lançamento do dado, caso o número de tentativas fosse cinco ($n = 5$), teríamos as probabilidades da Tabela 6 para cada número de “sucessos” *k*, dado que a probabilidade (*p*) de em um lançamento conseguir o número seis é de 1/6 (≅16.7%).

Tabela 6 - Probabilidades do exemplo de lançamento de dado.

k (Qtde. de "sucessos")	P(X = k)
0	40,2%
1	40,2%
2	16,1%
3	3,2%
4	0,3%
5	≅0,0%
Total	100,0%

No caso de VAs contínuas, *i.e.* aquelas em que seus contradomínios são intervalos ou conjuntos de intervalos em $\mathbb{R}$, é necessário introduzir o conceito de função densidade de probabilidade (PDF, em inglês). Não faria sentido associar probabilidades aos valores individuais da variável contínua, então associamos probabilidades a intervalos desta variável. A definição matemática da PDF, simbolizada por *f(x)*, onde *x* é um valor da variável, encontra-se nas Equações 24a, 24b e 24c [11].

$$f(x)\begin{cases}\geq 0, para\ x \in \mathbb{R}_x \\ = 0, para\ x \notin \mathbb{R}_x\end{cases} \tag{24a}$$

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1 \tag{24b}$$

$$P(a \leq x \leq b) = \int_a^b f(x)\,dx \tag{24c}$$

Graficamente, o conceito torna-se um pouco mais claro. A probabilidade é a integral (área) embaixo da curva *p(x)* fictícia para o intervalo de valores entre *a* e *b* da Figura 11.

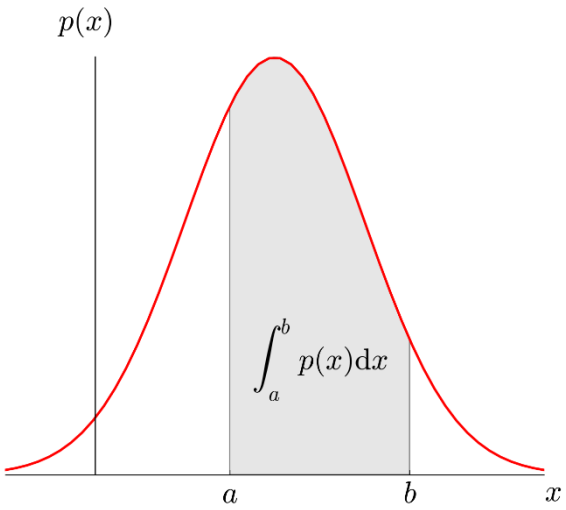

Figura 11 - Exemplo gráfico da definição de função densidade de probabilidade e do cálculo da probabilidade. Fonte: Adaptado de [12].

Dentre as diversas distribuições de probabilidade aplicáveis a VAs contínuas, a que merece maior destaque é sem dúvidas a distribuição Gaussiana, comumente conhecida como distribuição Normal. Sua PDF é dada pela Equação 25 [8],

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{25}$$

onde $-\infty < x < +\infty$, σ é o desvio padrão da variável com $\sigma > 0$ e μ é sua média aritmética com $-\infty < \mu < +\infty$. σ e μ devem ser constantes.

A distribuição Normal modela variáveis nos mais diversos campos da Ciência, desde Matemática e Economia até Anatomia, e por isso é tão importante. Suas variações, como a distribuição Normal Reduzida [quando $X \sim N(0,1)$] e a Log-Normal[3], também são bastante comuns. Outro fato que justifica o grande uso da distribuição Gaussiana é o Teorema do Limite Central, que postula que a soma de variáveis aleatórias independentes, discretas ou contínuas, tem distribuição aproximadamente Normal. Na prática, verifica-se empiricamente que essa aproximação é satisfatória para a soma de mais de vinte e cinco variáveis.

2.2.2. Valor Esperado, Variância e Covariância de Variáveis Aleatórias

O valor esperado [*E(X)*] de uma variável aleatória, também conhecido como expectância ou esperança matemática, é seu valor médio calculado com base em sua distribuição de probabilidades [11]. As Eq. 26 e 27 descrevem o conceito.

$$E(X) = \sum_i x_i \, P(x_i) \tag{26}$$

para variáveis discretas e

$$E(x) = \int_{-\infty}^{+\infty} x f(x) \, dx \tag{27}$$

para variáveis contínuas.

Fazendo uma analogia com a Física, o valor esperado de uma variável é seu "centro de massa". Observando a Eq. 27, essa analogia fica ainda mais clara.

A variância de uma VA é dada pela Equação 28. O desvio-padrão, assim como definido na seção 2.1.2, é a raiz quadrada positiva da variância.

[3] Utilizada, por exemplo, para modelar permeabilidades absolutas de reservatórios de petróleo. Se uma VA "*X*" tem distribuição Normal, então log(*X*) tem distribuição Log-Normal.

$$V(X) = E\{[X - E(X)]^2\} = E(X^2) - [E(X)]^2 \quad (28)$$

A covariância entre duas VAs (*X* e *Y*), ou seja, a medida de como as mesmas variam em conjunto, é dada pela Equação 29. É importante salientar que se *X* e *Y* são independentes, a covariância é nula, porém a recíproca não é verdadeira [11].

$$COV(X,Y) = E(X,Y) - E(X)E(Y) \quad (29)$$

Karl Pearson, proeminente Estatístico do século XX e fundador do Departamento de Estatística Aplicada na *University College of London*, desenvolveu o Coeficiente de Correlação de Pearson (CCP), que se utiliza da definição de covariância:

$$CCP = \frac{COV(X,Y)}{\sigma_X \sigma_Y}, \quad (30)$$

onde *COV(X,Y)* é a covariância de *X* e *Y* e σ_X, σ_Y são os desvios padrões de cada uma. CCP varia entre -1 (correlação linear negativa perfeita) e +1 (correlação linear positiva perfeita). Este coeficiente é a base para construção da matriz de correlação da seção 5 deste trabalho.

2.3. Estatística Indutiva

Também denominada “Inferência Estatística”, é o campo da Estatística de maior interesse para os desenvolvimentos e análises contidos na seção 5. Conforme já mencionado, seu grande objetivo é, através da Teoria da Probabilidade e suas ferramentas, construir modelos e inferências sobre populações através da análise de somente uma parte do todo, *i.e.* utilizando amostras [8]. Toma-se a parte (amostra) pelo todo (população), porém com certo intervalo de confiança, que funciona como uma “qualificação” do processo. Este conceito será elaborado ainda nesta seção. A Figura 12 ilustra o conceito de amostragem.

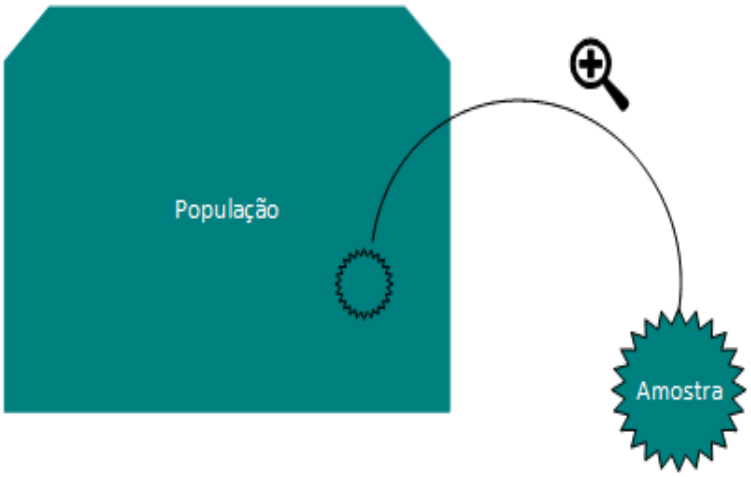

Figura 12 - O processo de amostragem em uma população.

É intuitivo pensar que o processo de amostragem é determinante na inferência estatística: uma amostra representativa de uma população fornecerá resultados confiáveis e precisos, enquanto que uma amostra ruim tem o poder arruinar conclusões e até mesmo induzir questionamentos sobre a credibilidade de um estudo indutivo.

2.3.1. Estatísticas, Estimadores e Distribuições Amostrais

Imagine o seguinte cenário: um testemunho (amostra cilíndrica de uma rocha) de 150 metros e sua população de porosidades (espaços vazios entre os grãos da rocha). Retiremos um *plug* (pequena amostra cilíndrica) deste testemunho (vide Figura 13). Assumindo que a porosidade varia ao longo do testemunho de acordo com uma distribuição Normal, os Y_i's (*i*-ésimos elementos da amostra) têm todos a mesma distribuição de probabilidade da população. Trata-se de uma amostragem sem reposição (não é como retirar bolas de bilhar de um recipiente e retorná-las), mas como a população é muito maior que a amostra (um *plug* tem dimensões de centímetros, enquanto um testemunho é da ordem de centenas de metros), os Y_i's são aproximadamente independentes. Então, podemos gerar estatísticas, *i.e.* qualquer valor calculado em função dos valores amostrais, desse *plug*.

Figura 13 - Um testemunho rochoso. As pequenas fendas circulares são *plugs* que foram retirados como amostras. Fonte: [13].

As Equações 31 e 32 descrevem duas estatísticas comuns: a média e variância amostrais, respectivamente.

$$\bar{Y} = \frac{Y_1 + Y_2 + Y_3 + \cdots + Y_n}{n} \tag{31}$$

$$S^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{X})^2}{n} \tag{32}$$

Já um estimador é uma estatística que tem o objetivo de inferir (estimar) o valor de um parâmetro populacional. É desejável que os estimadores sejam não-tendenciosos (Eq. 33a) e consistentes (Eq. 33b):

$$E(\hat{\theta}) = \theta\,; \tag{33a}$$

e

$$\lim_{n\to\infty} V(\hat{\theta}) = 0\,, \tag{33b}$$

onde $\hat{\theta}$ é o estimador de θ, um parâmetro populacional, e *n* é o tamanho da amostra.

A distribuição de probabilidade de uma estatística é intitulada distribuição amostral ou distribuição por amostragem. Pode-se provar que a distribuição amostral da variância, por exemplo, é qui-quadrado (χ^2) [14].

2.3.2. Estimação de Parâmetros

A estimação de parâmetros está centrada tanto em encontrar estimadores quando em avaliá-los. Em geral, são atividades relacionadas: muitas vezes os métodos de avaliação irão sugerir novos estimadores [8]. Há basicamente dois tipos de estimação: pontual e por intervalo.

Na estimação pontual, busca-se um estimador pontual, que é qualquer função de uma amostra. Portanto, toda estatística é um estimador pontual. Não é uma técnica muito utilizada devido ao grande erro inerente ao processo.

Na estimação por intervalo, busca-se um estimador que esteja dentre um intervalo, chamado de intervalo de confiança. Para a construção deste intervalo, defini-se um nível de significância α, *i.e.* a probabilidade de erro/desvio em relação ao real parâmetro da população. Na estimação de uma média populacional (μ), por exemplo, se a estatística $\bar{X}$ (média amostral) for o estimador, gostaríamos que este ficasse dentre do intervalo de confiança de forma que

$$P(\bar{X} - e_o < \mu < \bar{X} + e_o) = 1 - \alpha\,, \tag{34}$$

onde e_o é o valor que estipula o tamanho do intervalo de confiança e que pode ser calculado através de alguns passos algébricos [8]. O que a Equação 34 descreve é que a probabilidade de acerto, ou seja, de o estimador ficar dentro do intervalo, é de $1 - \alpha$.

Os níveis de significância mais comuns na literatura e na indústria como um todo são 5% e 10%, que geram, respectivamente, intervalos de confiança de 95% e 90% (a seção 2.3.3 detalha mais o assunto). Para ilustrar estes conceitos, considere a seguinte afirmação hipotética: “a probabilidade de encontrar petróleo na Bacia de Campos é em média 52%, com erro de três pontos percentuais para mais ou para menos”, ou seja, a cada 100 poços perfurados na Bacia de Campos, em média apenas 52 terão petróleo. Os três pontos percentuais são os limites da estimação por intervalo ($\pm e_o$).

É muito importante destacar que mesmo quando se calcula um intervalo de confiança, ainda assim não é possível afirmar se o real parâmetro populacional está contido ou não neste intervalo [15]. Tem-se confiança somente em que o estimador está contido no intervalo com determinada probabilidade. No exemplo supracitado, é possível que a real porcentagem de sucesso na descoberta de petróleo seja, por exemplo, 30%, *i.e.* um valor fora do intervalo construído (49% a 55%).

2.3.3. Testes de Hipóteses

Ao invés de encontrar um estimador para um parâmetro, é possível (e conveniente) hipotetizar um valor para o mesmo e depois utilizar informações da amostra para confirmar ou recusar esse valor. A esse processo se dá o nome de Teste de Hipóteses [11]. Nesta seção, a técnica será exemplicada utilizando um teste de hipótese unilateral[4] para a média de uma população que segue a distribuição Normal, porém, há testes bilaterais, testes para a variância, proporção, entre outros.

Consideremos o seguinte exemplo, adaptado de SALDANHA [9]: uma empresa que fabrica umbilicais[5] para campos do pré-sal brasileiro recebe mangueiras termoplásticas de poliuretano (um tipo de elastômero), necessárias para a confecção dos umbilicais, em grandes lotes. A carga média de ruptura à tração dessas mangueiras é de 350 kgf/cm² (μ_t). A empresa não se preocupa com a eventualidade de $\mu_t \geq 350$, mas quer proteger-se contra aceitar lotes com $\mu_t < 350$ por questões de segurança operacional *offshore*. Mais especificamente, quanto menor for o valor de μ_t, menor deve ser a probabilidade de aceitar o lote. Ao mesmo tempo, ela não quer rejeitar lotes com $\mu_t \geq 350$. Suponhamos que o desvio-padrão do lote (populacional) é conhecido, $\sigma = 6$ kgf/cm². A empresa, então, adotou o seguinte procedimento:

a) Escolher ao acaso *n* mangueiras do lote (amostragem);
b) Submetê-las aos ensaios mecânicos;
c) Calcular $\bar{x}$, a média das *n* cargas das mangueiras da amostra;
d) Rejeitar o lote caso $\bar{x} <$ C e, caso contrário, aceitá-lo.

[4] Também chamado de unicaudal.
[5] Conjugado de cabos (*bundle*) que supre equipamentos submarinos com fluidos hidráulicos, energia elétrica e eventualmente químicos.

Os objetivos da empresa são obter probabilidades baixas de rejeitar lotes bons e de aceitar lotes ruins. Porém, não sabe *a priori* quais valores de *n* e *C* deverão ser escolhidos. Sem muito critério, como parâmetros iniciais, foram escolhidos $n = 25$ e $C = 345$ kgf/cm².

O que a empresa está testando é a chamada Hipótese Nula (H_0), que é a hipótese de que o lote é bom ($\mu_t \geq 350$), contra uma Hipótese Alternativa (H_1), de que o lote é ruim ($\mu_t < 350$). A decisão será, então, aceitar ou rejeitar H_0 (vide Tabela 7).

Tabela 7 - O Teste de Hipóteses e as possíveis configurações.

		Realidade	
		H_0 verdadeira ($\mu_t \geq 350$)	H_0 falsa ($\mu_t < 350$)
Decisão	Aceitar H_0 ($\bar{x} \geq 345$)	OK	Erro do tipo II**
	Rejeitar H_0 ($\bar{x} < 345$)	Erro do tipo I*	OK

*Rejeitar H_0 quando ela é verdadeira.
**Aceitar H_0 quando ela é falsa.

Aceitar H_0 é um evento aleatório, cuja probabilidade depende do valor verdadeiro de μ_t. Explicita-se isso como

$$P(Aceitar\ H_0) = L(\mu_t)\,, \tag{35}$$

ou ainda

$$P(\mu_t \geq 350) = L(\mu_t)\,, \tag{36}$$

onde $L(\mu_t)$ é a função característica de operação do teste, também chamada de curva característica de operação (CCO). Para a construção dessa curva, assume-se que a população tem distribuição Normal (portanto, as amostras e a média amostral também):

$$\bar{x} \sim N\left(\mu_t, \frac{\sigma^2}{n}\right) \rightarrow \bar{x} \sim N(\mu_t, 1{,}44) \tag{37}$$

$$L(\mu_t) = P(\mu_t \geq 350) = P\left(\frac{\bar{x} - \mu_t}{1{,}44} \geq \frac{345 - \mu_t}{1{,}44}\right) = P\left(z \geq \frac{345 - \mu_t}{1{,}44}\right) \qquad (38)$$

onde z é a variável reduzida da Normal. Para diferentes μ_t's e valores críticos (C), podemos construir a Tabela 8 e a Figura 14.

Tabela 8 - Função característica de operação com variação nos valores críticos.

	L(μt)		
μt	C = 345	C = 346	C = 347
343	0,082433	0,018610	0,002737
343,5	0,148783	0,041272	0,007538
344	0,243702	0,082433	0,018610
344,5	0,364212	0,148783	0,041272
345	0,500000	0,243702	0,082433
345,5	0,635788	0,364212	0,148783
346	0,756298	0,500000	0,243702
346,5	0,851217	0,635788	0,364212
347	0,917567	0,756298	0,500000
347,5	0,958728	0,851217	0,635788
348	0,981390	0,917567	0,756298
348,5	0,992462	0,958728	0,851217
349	0,997263	0,981390	0,917567
349.5	0,999111	0,992462	0,958728
350	0,999742	0,997263	0,981390

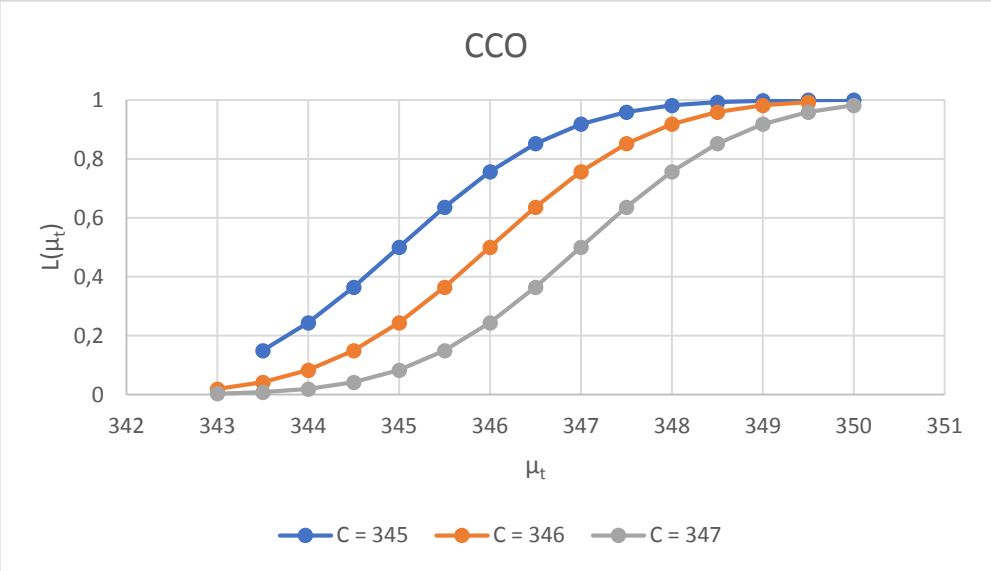

Figura 14 - Variação da CCO de acordo com os valores críticos.

Observe que o aumento no valor crítico desloca a CCO para a direita (Figura 14). Isso significa que há um aumento na probabilidade de erro do tipo I, ou seja, de rejeitar um lote bom. É intuitivo pensar que isso ocorre, já que o aumento de *C* significa maior conservadorismo. Observe na Tabela 8 que enquanto para *C* = 345 a probabilidade de erro era de ≅ 0,03% [1 – P(μ_t = 350)], para *C* = 347 a probabilidade é de ≅ 1,86%. Ao mesmo tempo, há uma diminuição da probabilidade de erro do tipo II, o que leva à seguinte conclusão: a única maneira de reduzir os dois erros conjuntamente é aumentar *n*, o número de constituintes da amostra. Na CCO, isso se refletiria em um aumento de inclinação das curvas.

Além da CCO, outros conceitos fundamentais da teoria de Testes de Hipóteses são:

- Nível de significância: geralmente simbolizado por *α*, segue a mesma definição da seção 2.3.2. É sinônimo da probabilidade de se rejeitar a hipótese nula quando ela é verdadeira (erro do tipo I). Quando maior for *α*, mais confiança se tem na hipótese nula [9].
- Significância estatística: conceito primeiramente instituído por Sir Ronald Fisher [16], é, em termos práticos, a medida de validade do resultado de um teste de hipótese.

Para o cálculo de significância estatística, que é um procedimento de teste de hipótese, são utilizadas estatísticas de teste. Geralmente utiliza-se a estatística de teste *t* e o Teste t de Student[6], também conhecido como “Teste t”, especialmente nos casos em que não se conhece o desvio padrão populacional (*σ*), somente o amostral (*s*). A estatítica *t* segue uma distribuição t de Student, que é similar à Normal, porém com uma variância maior. A Eq. 39 descreve sua formulação [9]:

$$t_{n-1} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}\ , \tag{39}$$

onde μ_0 é o valor para a média hipotetizado em H_0 (no exemplo da empresa fabricante de umbilicais, μ_0 = 350) e *s* é o desvio padrão amostral.

[6] Student é o pseudônimo de William Sealy Gosset, um químico da cervejaria irlandesa Guinness que desenvolveu um teste de hipóteses como forma de controlar a qualidade da cerveja de tipo *stout* com um custo baixo. O pseudônimo foi escolhido pois a cervejaria acreditava se tratar de um segredo industrial [5].

No exemplo da empresa de umbilicais, o desvio padrão populacional é conhecido *a priori*. Neste caso, podemos utilizar a estatística de teste da Eq. 40, que segue uma distribuição Normal Reduzida [z ~ N(0,1)].

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \tag{40}$$

Suponhamos que, considerando H_0 verdadeiro, a empresa gostaria de saber qual a probabilidade de a carga média de ruptura à tração ($\bar{x}$) do lote ser menor ou igual a 348 kgf/cm². Utilizando a estatística de teste z, temos que

$$P(\bar{x} \leq 348) = P\left(\frac{\bar{x} - 350}{6/5} \leq \frac{348 - 350}{6/5}\right) = P(z \leq -2) \cong 4{,}78\% \,. \tag{41}$$

O resultado da Eq. 41 é chamado de valor-p (também denominado probabilidade de significância), que é interpretado como o menor nível de significância com que se rejeita a hipótese nula. Um valor-p pequeno significa que a probabilidade de se obter um valor da estatística de teste como o calculado é muito improvável, o que leva à rejeição da hipótese nula. Se o nível de significância estiver acima do valor-p, não se rejeita H_0. Considerando um nível de significância de 5% para o exemplo, não rejeitaríamos a hipótese nula, o que neste caso significa que o lote de mangueiras termoplásticas encontra-se dentro da especificação de carga de ruptura à tração.

Apesar de ser contraintuitivo, o maior objetivo de um teste de hipóteses é rejeitar a hipótese nula, que é construída de forma a representar uma situação do tipo “padrão”, *i.e.* comum ou esperada.

2.3.4. ANOVA

ANOVA é o acrônimo de *Analysis of Variance*, ou Análise de Variância na tradução livre. Apesar do nome, a técnica na verdade utiliza as técnicas de testes de hipóteses para analisar variações entre médias, não entre variâncias. É uma das ferramentas estatísticas mais utilizadas na Academia e na Indústria no geral [8].

Há mais de uma classe de ANOVA, como por exemplo a MANOVA (*Multivariate Analysis of Variance*), aplicada no caso de existir mais de uma variável dependente, e a ANOVA fatorial, na qual é possível analizar efeitos interativos de maior ordem entre várias variáveis independentes [17]. O desenvolvimento das definições deste capítulo será feito com base na *One-Way ANOVA*, aplicável a casos com somente uma variável independente, visto que a fundamentação matemática de todos os tipos de ANOVA têm como base o caso mais simples.

Considere a seguinte situação hipotética: uma empresa de Geofísica[7] gostaria de estudar se sua demanda por campanhas sísmicas[8] 4-D, uma tecnologia de última geração, é sazonal. Para isso, compilou dados históricos dos cinco últimos anos, gerando sua população de dados. É possível dividí-la em *k* populações (ou subpopulações) de acordo com alguns critérios, chamados de tratamentos, e amostrar alguns valores delas. No caso, o tratamento escolhido foram os meses de um ano. Então, buscou-se realizar uma *One-Way* ANOVA.

O teste de hipóteses fundamental de uma ANOVA é o descrito pelas Equações 42a e 42b:

$$H_0\colon\ \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k\ , \tag{42a}$$

$$H_1\colon\ \mu_i \neq \mu_j\ para\ algum\ i,j\ , \tag{42b}$$

onde μ_i é a média da população *i*, *i* = 1, 2, 3, ... , *k* e *k* é o número de populações.

Para a aplicação de uma ANOVA, as seguintes condições devem ser satisfeitas [9]:

a) As *k* populações têm a mesma variância (condição de homocedasticidade) ou aproximadamente a mesma. É a condição mais importante para a realização da ANOVA [8];
b) As *k* populações têm distribuição Normal ou com a curtose próxima a da Normal (mesocúrtica);

[7] Ciência que busca obter informações e modelos da superfície e subsuperfície terrestres a partir de métodos acústicos e eletromagnéticos.
[8] Operações que envolvem métodos acústicos e que buscam realizar um mapeamento das camadas de rocha em subsuperfície através da diferença de tempos de reflexão das ondas sísmicas.

c) As amostras das *k* populações são independentes.

Antes do desenvolvimento matemático, faz-se necessário definir algumas notações:

- Indexadores: *j* = 1, 2, ... , *n* (número de elementos de uma amostra) e *i* = 1, 2, ... , *k* (número de subpopulações);
- x_{ij}: *j*-ésimo valor da *i*-ésima subpopulação;
- $T_i = \sum_j x_{ij}$, ou seja, a soma dos valores da *i*-ésima subpopulação;
- $Q_i = \sum_j x_{ij}^2$, ou seja, a soma dos quadrados dos valores da *i*-ésima subpopulação;
- $\overline{x_i} = \frac{T_i}{n}$, ou seja, a média dos valores da *i*-ésima subpopulação;
- $T = \sum_i T_i$, ou seja, a soma dos valores de todas as subpopulações;
- $Q = \sum_i Q_i$, ou seja, a soma dos quadrados dos valores de todas as subpopulações;
- $\bar{\bar{x}} = \frac{T}{nk}$, ou seja, a média dos *nk* valores.

Quando H_0 for verdadeira, haverá três estimadores distintos da variância populacional σ^2, comum às *k* subpopulações:

1. Estimativa total, dada pela Eq. 43.

$$S_T^2 = \frac{\sum_i \sum_j \left(x_{ij} - \bar{\bar{x}}\right)^2}{nk - 1} = \frac{Q - \frac{T^2}{nk}}{nk - 1} = \frac{SQT}{nk - 1}\,, \tag{43}$$

onde *SQT* é a soma de quadrados total e S_T é a variância amostral total. Se H_0 for falsa, há superestimação da variância populacional.

2. Estimativa entre amostras, dada pela Eq. 44.

$$S_E^2 = \frac{n \sum_i (\overline{x_i} - \bar{\bar{x}})^2}{k - 1} = \frac{\frac{\sum_i T_i^2}{n} - \frac{T^2}{nk}}{k - 1} = \frac{SQE}{k - 1}\,, \tag{44}$$

onde *SQE* é a soma de quadrados entre amostras. Se H_0 for falsa, também há superestimação da variância populacional.

3. Estimativa residual, dada pela Eq. 45.

$$S_R^2 = \frac{\sum_{i=1}^{k} \frac{\sum_j \left(x_{ij} - \bar{x}_i\right)^2}{n-1}}{k} = \frac{Q - \frac{\sum_i T_i^2}{n}}{k(n-1)} = \frac{SQR}{k(n-1)}\,, \tag{45}$$

onde *SQR* é a soma de quadrados residual, dentre amostras. Se H_0 for falsa, não há sobrestimação ou subestimação da variância populacional, pois sua definição baseia-se somente em valores das amostras.

Note que:

$$SQT = SQE + SQR\,, \tag{46}$$

$$\frac{SQT}{\sigma^2} = \frac{SQE}{\sigma^2} + \frac{SQR}{\sigma^2}\,. \tag{47}$$

Todos os termos da Eq. 47 têm distribuição de qui-quadrado com, respectivamente, *nk*-1, *k*-1 e *k*(*n*-1) graus de liberdade [9]. Então:

$$\chi^2_{nk-1} = \chi^2_{k-1} + \chi^2_{k(n-1)}\,. \tag{48}$$

Finalmente, quando H_0 for verdadeira, teremos:

$$F = \frac{S_E^2}{S_R^2}\,, \tag{49}$$

onde *F* é a estatística de teste e tem distribuição F de Snedecor[9] [18]. Se H_0 for falsa, F tenderá a crescer [9]. Rejeita-se H_0 quando:

$$F > F_{k-1;k(n-1);\alpha}\,, \tag{50}$$

[9] Snedecor denominou-a "F" como homenagem a Fisher. Define-se para uma variável aleatória *X* que tem distribuição F de Snedecor que *X* ~ F(m,n), onde *m* são os graus de liberdade no denominador e *n* são os graus de liberdade no numerador. A notação $F_{m,n}$ também é utilizada.

onde o termo da direita é tabelado (vide Anexo II) e por vezes chamado de F crítico e α é o nível de significância adotado.

No exemplo da empresa de Geofísica, podemos organizar os dados conforme a Tabela 9.

Tabela 9 - Conjunto de dados históricos de demanda por campanhas sísmicas.

Mês	Demanda por Campanhas Sísmicas (nº de contratos)					Média	Variância
	Ano 1	Ano 2	Ano 3	Ano 4	Ano 5		
Janeiro	27	12	44	6	11	20,0	241,5
Fevereiro	21	6	33	19	4	16,6	141,3
Março	2	20	48	28	22	24,0	274,0
Abril	21	21	9	28	35	22,8	93,2
Maio	26	19	44	39	17	29,0	144,5
Junho	21	18	25	5	54	24,6	326,3
Julho	11	59	30	30	31	32,2	294,7
Agosto	30	15	0	22	9	15,2	133,7
Setembro	51	20	10	30	23	26,8	234,7
Outubro	34	2	54	17	13	24,0	413,5
Novembro	28	1	17	10	8	12,8	104,7
Dezembro	22	9	38	2	7	15,6	211,3

Neste exemplo, a população são as demandas por campanhas sísmicas e as subpopulações são as demandas por mês (k = 12), visto que o tratamento (critério) são os meses do ano. As amostras das subpopulações são os valores históricos dos últimos cinco anos (n = 5). Há 11 graus de liberdade no numerador (k – 1) e 48 no denominador [$k(n-1)$].

Note na Tabela 9 que as variâncias são aproximadamente similares na maioria dos casos (entre 100 e 400, na mesma ordem de grandeza). Portanto, uma das premissas para aplicação da ANOVA foi satisfeita.

Para verificar se as populações têm distribuição normal, podemos utilizar o Gráfico de Probabilidade Normal, conhecido como NPP (*Normal Probability Plot*). A reta vermelha na Figura 15 representa um ajuste perfeito à distribuição Normal. Poderíamos usar também algum Teste de Aderência[10], como o Teste de Qui-Quadrado [19] ou o de Kolmogorov-Smirnov [20].

[10] Testes que têm como objetivo verificar o ajuste de uma distribuição de probabilidade a uma população.

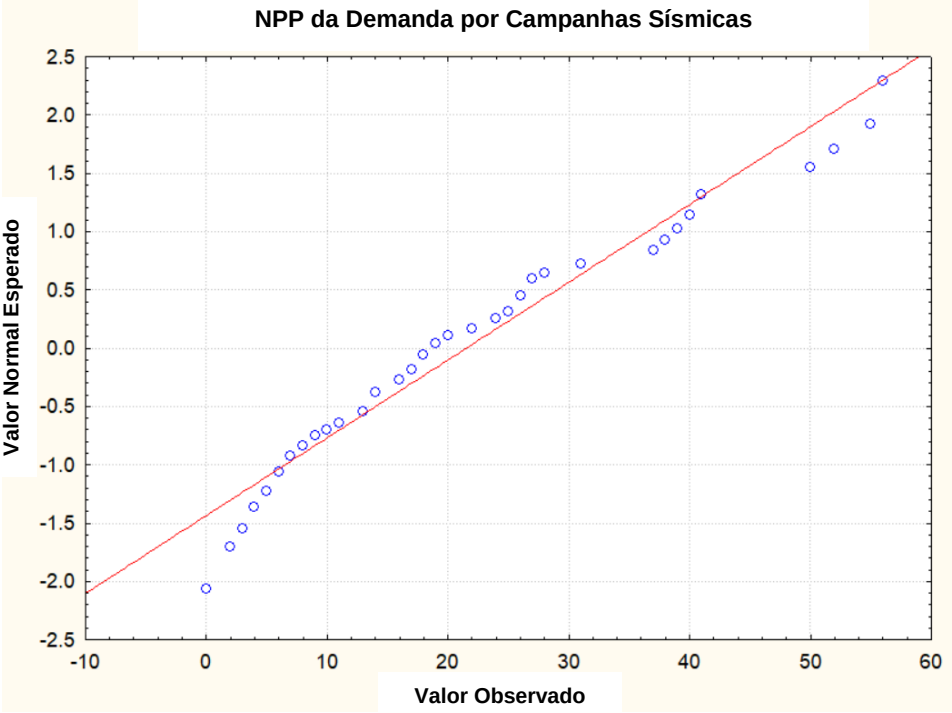

Figura 15 - NPP gerado no *software Statistica 7.0*® a partir dos dados de demanda por campanhas sísmicas (população). Supõe-se que sigam uma distribuição Normal para aplicação da ANOVA. Através de uma análise rápida do gráfico, é possível observar que a distribuição de probabilidades é aproximadamente Normal.

Percebe-se pela Figura 15 de que a premissa de normalidade foi atingida.

Quanto à apresentação dos resultados, a "Tabela ANOVA" é a forma mais comum. Ela contém colunas com os graus de liberdade (DOF, *Degrees of Freedom*), Soma dos Quadrados (SS, *Sum of Squares*), *i.e. SQE* e *SQR*, e Média da Soma dos Quadrados (MS, *Mean Square*), *i.e SQE*/*k*-1 e *SQR*/*k*(*n*-1), calculadas para variações entre os grupos de tratamentos e dentre grupos de tratamentos (residual).

Então, aplicando a ANOVA com um nível de significância de 5% (α = 5%), obtemos o resultado da Tabela 10 (o Anexo I contém um passo a passo realizado no *software Statistica 7.0*® para a geração do resultado). Note na tabela que o total de graus de liberdade é o número total de valores (N = 60) menos 1 (N – 1).

Tabela 10 - Tabela ANOVA para o exemplo.

Fonte de Variação	SS	DOF	MS	F	valor-p
Mês (entre tratamentos)	2028.6	11	184.41818	0.885739	0.560094
Residual (dentre tratamentos)	9994	48	208.20833		
Total	12022,6	59	392,62651		

Como o valor da estatística F não é maior que o valor crítico (caso fosse, o *software* marcaria o valor em vermelho), o que é equivalente a afirmar que o valor-p do resultado está acima no nível de significância adotado (vide Tabela 10), não se rejeita a hipótese nula considerando 5% de significância. Portanto, segundo a ANOVA, a demanda da empresa de Geofísica não é sazonal, *i.e.* não é influenciada pelo mês do ano.

2.4. Regressão Linear

2.4.1. Regressão Linear Simples

Diversas vezes, em um diagrama de dispersão[11], os pontos obtidos por análises experimentais sugerem a existência de uma relação funcional entre as duas variáveis do gráfico. Surge então o problema de se determinar uma função que descreva esse relacionamento. Sir Fisher denominou-o como "regressão" [4]. Se a função obtida pela regressão for conhecida, o problema é relativamente simples e fica reduzido à estimação de parâmetros. Porém, se não for conhecida *a priori*, ela deverá ser inferida juntamente aos seus parâmetros.

No caso em que a curva de regressão é uma reta, a técnica é intitulada Regressão Linear Simples. O termo "simples" indica que o problema possui somente duas variáveis. Para mais de duas variáveis, é necessário realizar uma Regressão Linear Múltipla [4].

Considere um problema que envolve duas variáveis, *x* e *y*. A relação entre elas pode ser determinística, *i.e.* cada valor de *x* corresponde a um e somente um valor de *y*, ou estocástica, *i.e.* cada valor de *x* corresponde a uma distribuição de probabilidade de *y*. As Equação 51 e 52 descrevem as relações:

$$y = f(x)\,, \tag{51}$$

[11] Gráfico que utilizada coordenadas cartesianas para exibir valores de um conjunto de dados.

no caso determinístico, e

$$y = f(x) + \zeta\,, \tag{52}$$

no caso estocástico. ζ é uma variável aleatória e, então, *y* também é.

Tomando a relação estocástica, pode-se construir um modelo trivial de regressão linear simples, conforme a Eq. 53:

$$y_i = A + Bx_i + \zeta_i \;\; para\; i = 1, 2, \dots, n\,; \tag{53}$$

onde *y* é a variável dependente (aleatória), y_i é o valor de *y* em uma observação/medição *i*, *x* é a variável independente (não aleatória e também chamada de fator ou variável explicativa), x_i é o valor de *x* em uma medição *i*, ζ é uma perturbação aleatória, ζ_i é o valor de ζ em uma medição *i* e *A*, *B* são parâmetros da regressão, *a priori* desconhecidos.

Neste modelo, ζ tem o objetivo de sintetizar toda a aleatoriedade de um fenômeno, desde as incertezas relacionas ao fenômeno em si até os erros experimentais envolvidos. Além disso, o modelo tem as seguintes hipóteses:

a) A variável *x* e o conjunto de x_i's são não-aleatórios;
b) As variáveis aleatórias ζ_i's são independentes (covariância nula), $E(\zeta_i) = 0$ e $V(\zeta_i) = \sigma^2$, *i.e.* a variância da população.
c) $\zeta_i \sim N(0,\sigma^2)$, ou seja, as perturbações aleatórias têm distribuição Normal com média igual a zero e variância igual a σ^2.

Então,

$$E(y_i) = A + Bx_i + \cancelto{0}{E(\zeta_i)} = A + Bx_i\,. \tag{54}$$

A Eq. 54 é intitulada reta de regressão populacional. A Figura 16 ilustra os conceitos.

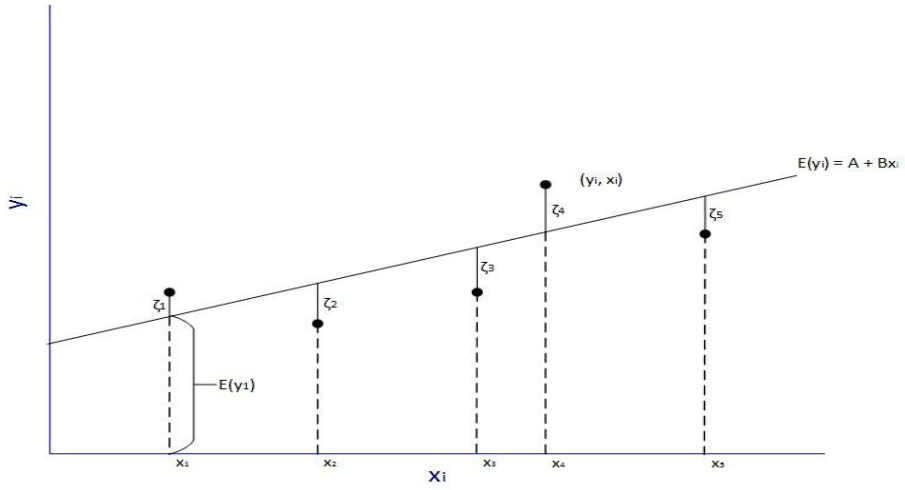

Figura 16 - Representação gráfica do modelo de Regressão Linear Simples. Observe a reta de regressão populacional.

Quando estimamos *A* e *B* (por qualquer método), obtemos a chamada reta de regressão amostral, dada por

$$\widehat{y_i} = a + bx_i, \tag{55}$$

onde $a = \hat{A}$, $b = \hat{B}$ (estimadores) e $\widehat{y_i}$ é o valor de y_i estimado para x_i.

Os valores de y_i e $\widehat{y_i}$ geralmente não são iguais. A diferença entre eles é chamada resíduo (Eq. 56), que pode ser entendido como um estimador da perturbação aleatória ζ_i .

$$e_i = y_i - \widehat{y_i} \tag{56}$$

Logo,

$$y_i = a + bx_i \pm e_i = \widehat{y_i} \pm e_i \tag{57}$$

Há diversos métodos para estimar os parâmetros da regressão, como os que buscam o estimador linear de variância mínima (ou MELNT, melhor estimador linear não-tendencioso) ou o estimador de máxima verossimilhança (EMV). Porém, o mais comum dos métodos, aplicado massivamente na Academia e na Indústria, é o Método

dos Mínimos Quadrados, que busca encontrar os valores de *a* e *b* que minimizam a soma dos quadrados dos resíduos, isto é, a função:

$$S_{MMQ} = \sum_{i=1}^{n} {e_i}^2 = \sum_{i=1}^{n} (y_i - \hat{y_i})^2 = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \,. \tag{58}$$

Derivando parcialmente S_{MMQ} em relação a *a* e *b*, temos:

$$\frac{\partial S_{MMQ}}{\partial a} = \sum_{i} 2(y_i - a - bx_i) * (-1) \,; \tag{59}$$

$$\frac{\partial S_{MMQ}}{\partial b} = \sum_{i} 2(y_i - a - bx_i) * (-x_i) = 0 \,. \tag{60}$$

Então, igualando cada uma das derivadas a zero para encontrar o mínimo da função:

$$\frac{\partial S_{MMQ}}{\partial a} = 0 \;\rightarrow \sum_{i} y_i = na + b \sum_{i} x_i \,; \tag{61}$$

$$\frac{\partial S_{MMQ}}{\partial a} = 0 \;\rightarrow \sum_{i} x_i y_i = a \sum_{i} x_i + b \sum_{i} x_i^2 \,. \tag{62}$$

Após algumas manipulações algébricas obtém-se os estimadores das Equações 63 e 64.

$$a = \bar{y} - b\bar{x} \tag{63}$$

$$b = \frac{\sum_i x_i y_i}{\sum_i x_i^2} . \tag{64}$$

Após a regressão, os pontos que fogem da reta ajustada e da maioria dos dados experimentais são chamados de *outliers*. Seus valores residuais são altos e geralmente representam dados com grande erro experimental.

É importante destacar que antes de tentar se ajustar uma reta a dados experimentais, é importante verificar se há relação entre as variáveis de interesse. Isso não significa necessariamente que uma variável causa a outra, mas sim que há alguma associação significativa entre elas [21]. Uma das formas de se verificar essa associação é através do coeficiente de correlação de Pearson (CCP) (vide seção 2.2.2). Se não há nenhuma relação, a ajuste de uma reta aos dados provavelmente não fornecerá resultados úteis.

2.4.2. Regressão Linear Múltipla

Conforme já foi discutido, a regressão linear múltipla surge da tentativa de ajustar um modelo à relação entre duas ou mais variáveis explicativas e uma variável dependente através de uma reta. A reta de regressão populacional para *n* variáveis independentes (x_i's) é dada pela Eq. 65:

$$E(y_i) = A_0 + A_1 x_1 + A_2 x_2 + \cdots + A_n x_n \,, \tag{65}$$

A reta de regressão amostral para *k* observações das variáveis independentes é dada pela Eq. 66:

$$y_i = a_0 + a_1 x_{1j} + a_2 x_{2j} + a_n x_{nk} \pm e_j \,, \tag{66}$$

onde j = 1, 2, ... , k e e_j é o resíduo da observação j, cuja definição é a mesma da seção 2.5.1.

Utilizando o Método dos Mínimos Quadrados em um *software* como o *Statistica 7.0*®, facilmente os coeficientes a (estimadores de A) são obtidos.

Após a regressão, geralmente se realiza um teste t de Student para verificação de significância estatística dos parâmetros encontrados e para a construção de um intervalo de confiança para os mesmos a partir do nível de significância.

2.4.3. ANOVA utilizando Regressão

É possível realizar uma ANOVA (seção 2.3.4) via Regressão Linear, seja ela Simples ou Múltipla. No caso da Múltipla, há um ajuste dos graus de liberdade de forma que reflitam o número de variáveis intependentes [21]. O desenvolvimento desta seção será baseado no caso da Regressão Múltipla.

A ideia básica de uma regressão, ou seja, "Dados Experimentais = Modelo + Resíduos", pode ser reescrita da seguinte forma [21]:

$$(y_i - \bar{y}) = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)\,. \tag{67}$$

O termo à esquerda da Eq. 67 é a variação da observação *i* da variável dependente *y* em relação à média. O primeiro termo à direita é a variação do estimador de y_i em relação à média e o segundo termo é o resíduo. Elevando ambos os lados da equação ao quadrado e somando para *k* observações da variável, temos a Eq. 68:

$$\sum_{i=1}^{k} (y_i - \bar{y})^2 = \sum_{i=1}^{k} (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{k} (y_i - \hat{y}_i)^2\,. \tag{68}$$

Perceba que os termos da Equação 68 são equivalentes ao *SQT*, *SQE* e *SQR* da ANOVA (vide seção 2.3.4)! Podemos, então, definir uma tabela ANOVA via Regressão Linear Múltipla (Tabela 11).

Tabela 11 - Tabela ANOVA utilizando Regressão Linear Múltipla.

Fonte de Variação	SS	DOF	MS	F
Modelo	$\sum_{i=1}^{k}(\hat{y}_i - \bar{y})^2$	*n*	$\frac{\sum_{i=1}^{k}(\hat{y}_i - \bar{y})^2}{n}$	$\frac{\frac{\sum_{i=1}^{k}(\hat{y}_i - \bar{y})^2}{n}}{\frac{\sum_{i=1}^{k}(y_i - \hat{y}_i)^2}{k-n-1}}$
Residual	$\sum_{i=1}^{k}(y_i - \hat{y}_i)^2$	*k* - *n* - 1	$\frac{\sum_{i=1}^{k}(y_i - \hat{y}_i)^2}{k-n-1}$	
Total	$\sum_{i=1}^{k}(y_i - \bar{y})^2$	*k* - 1	$\frac{\sum_{i=1}^{k}(y_i - \bar{y})^2}{k-1}$	

A estatística de teste F tem distribuição F de Snedecor com *n* graus de liberdade no numerador e *k* - *n* - 1 no denominador. As hipóteses da ANOVA são:

$$H_0 : A_1 = A_2 = \cdots = 0 \tag{69}$$

$$H_1 : A_i \neq 0, para\ i = 1, 2, \dots, n \tag{70}$$

Finalmente, pode-se definir uma medida de qualidade de ajuste à regressão. A mais utilizada é o coeficiente de determinação, conhecido como R^2, que varia de 0 a 1 [4]. No caso de regressão múltipla, o mesmo é denominado coeficiente de determinação múltiplo e é dado pela Eq. 71:

$$R^2 = \frac{\frac{\sum_{i=1}^{k}(\hat{y}_i - \bar{y})^2}{n}}{\frac{\sum_{i=1}^{k}(y_i - \bar{y})^2}{k-1}} \tag{71}$$

Tanto para a regressão simples quanto para a múltipla, a interpretação prática do R^2 é o quanto da variabilidade dos dados experimentais é explicada pelas variáveis do modelo de regressão ajustado [21].

2.5. Diagrama de Pareto

A denominação “Diagrama de Pareto” pode ser tomada para mais de um tipo de análise. Seja qual for o caso, o título é uma homenagem ao Economista italiano do final do século XIX Vilfredo Pareto, conhecido pelo “Princípio de Pareto” (ou ainda “Regra 80/20”), o qual postula que para diversos eventos, aproximadamente 80% dos efeitos são oriundos de apenas 20% das causas [22]. Geralmente o nome é mais utilizado na indústria para os gráficos de Controle de Qualidade e Six Sigma[12] [23], os quais são histogramas combinados com uma curva de frequência acumulada (Fig. 17).

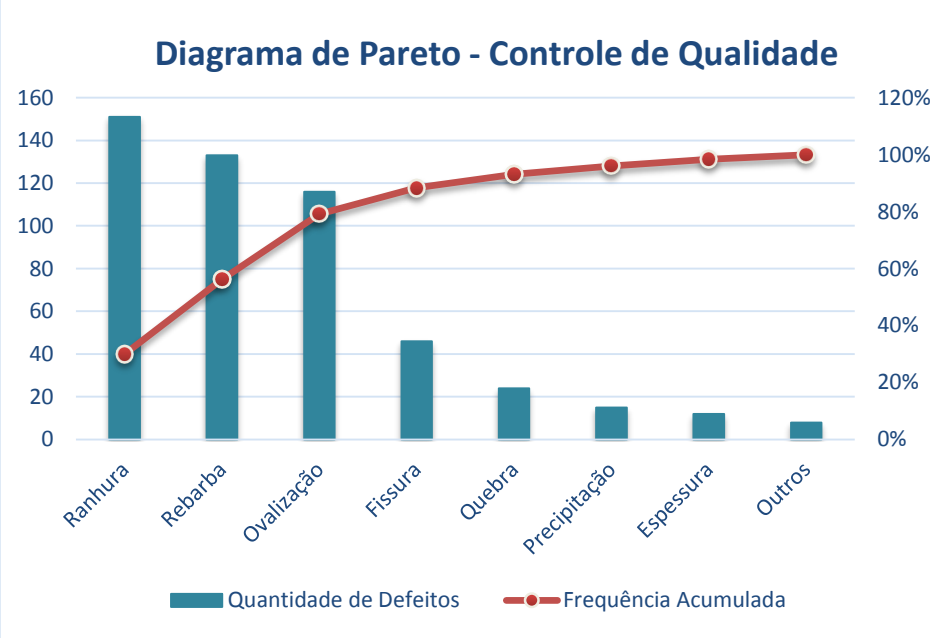

Figura 17 - Exemplo (fictício) de Diagrama de Pareto para o controle de qualidade em uma fábrica de dutos, relacionando defeitos diversos com a quantidade observada de cada um.

O outro uso comum ocorre em gráficos resultantes de análises da Estatística Indutiva, principalmente em técnicas de Planejamento Experimental[13]. Neste caso, o gráfico, conhecido também como “Diagrama de Efeitos de Pareto”, contém barras horizontais que estão associadas aos valores-t, a estatística do teste t de Student, das variáveis independentes e também uma linha-base que demonstra o ponto de valor-p

[12] Conjunto de práticas para eliminação de erros e desvios em processos industriais baseada na metodologia DMAIC (*Define, Measure, Analyze, Improve, Control*).
[13] Conhecido também como DOE (*Design of Experiments*), o termo se refere a um conjunto de técnicas, criadas originalmente por Sir Ronald Fisher, utilizadas para determinar os efeitos de variáveis de um experimento em seus resultados, constituindo, portanto, uma Análise de Sensibilidade avançada [24].

igual ao nível de significância adotado (Fig. 18). O tamanho das barras é proporcional à influência das variáveis independentes (“causas”) na variável dependente (“resultado”) estudada: quanto maior, mais influente é a variável (*i.e.* mais evidências se tem para rejeitar a hipótese nula que, relembrando, é descrita de forma a exprimir uma situação “padrão”, onde não há influências da variável independente na variável dependente). Além disso, caso o valor-t de uma variável seja maior que o $p = \alpha$, *i.e.* caso a barra ultrapasse a linha-base, há significância estatística, ou seja, há a rejeição da hipótese nula (H_0) de que a variável em questão não influi no resultado.

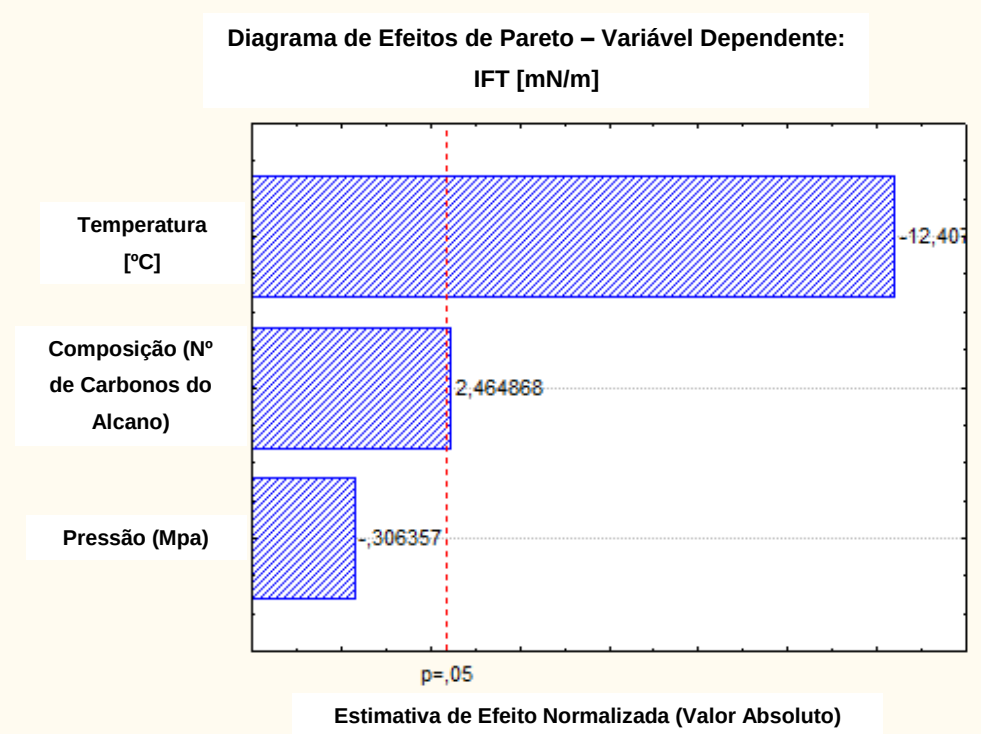

Figura 18 - Exemplo de um Diagrama de Efeitos de Pareto para um estudo de tensão interfacial (IFT) alcanos-água considerando três variáveis independentes: temperatura, composição do alcano e pressão. Note que a temperatura é a variável mais influente no resultado de IFT, com significância estatística.

3. Tópicos em Fenômenos de Superfície

3.1. Tensão Interfacial

Ao observar num dia chuvoso uma gota de chuva caindo do céu, percebe-se que seu formato aproxima-se de uma esfera, porém distorcida. Isto se deve aos efeitos de superfície em líquidos: há a tendência de redução da área superficial de forma que o maior número de moléculas esteja no corpo do líquido (*bulk*) e não na superfície, conforme será detalhado adiante. Como a esfera é a configuração geométrica e espacial que fornece a menor razão área/volume, é esta a forma que a gota assume ou tentará assumir. A distorção da gota é resultado da força de gravidade terrestre e da resistência do ar (vide Figura 19).

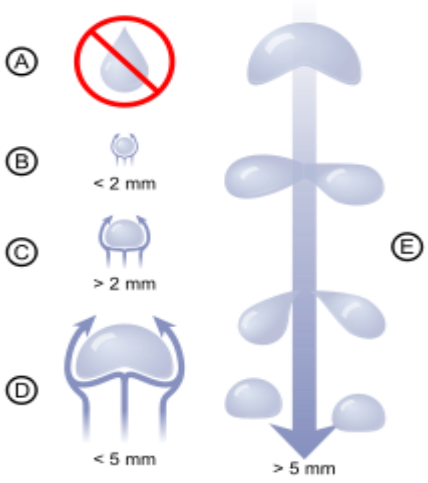

Figura 19 - Comparação entre gotas d'água muito pequenas (B) e muito grandes (D). Fonte: [25].

O que ocorre na formação de uma gota, na verdade, é uma tentativa de redução energética, que leva a uma redução de área superficial, como demonstram as Eq. 72a e 72b [26]. Assim como tudo na natureza tende a tomar o caminho de menor energia, não é diferente com uma gota d'água ou com qualquer outra gota. No caso das gotas de chuva, a diferença na resistência do ar, que é maior na base das gotas do que em seus topos, faz com que seus formatos se tornem achatados (vide "D" na Figura 19). Contudo, gotas de chuva muito pequenas são praticamente esféricas, visto que o volume e área das mesmas não são suficientes para causar diferenças significativas na

resistência do ar. Gotas muito grandes tornam-se tão achatadas que se dividem em gotas menores (item “E” da Figura 19).

$$dA = \gamma ds \tag{72a}$$

$$\left(\frac{\partial A}{\partial s}\right)_{P,T} = \gamma \tag{72b}$$

onde A é a energia livre de Helmholtz, γ é a tensão interfacial e *ds* é o infinitésimo de área superficial.

Consideremos dois fluidos imiscíveis, A e B (água e óleo, por exemplo). É intuitivo imaginar que uma interface entre eles irá se formar, separando-os e criando uma “zona de transição” entre eles. Isso ocorre pela diferença entre as forças intermoleculares com as moléculas quimicamente distintas em cada interface [27]: as moléculas localizadas na interface possuem energias potenciais maiores que as moléculas no corpo (interior) dos líquidos (Figura 20).

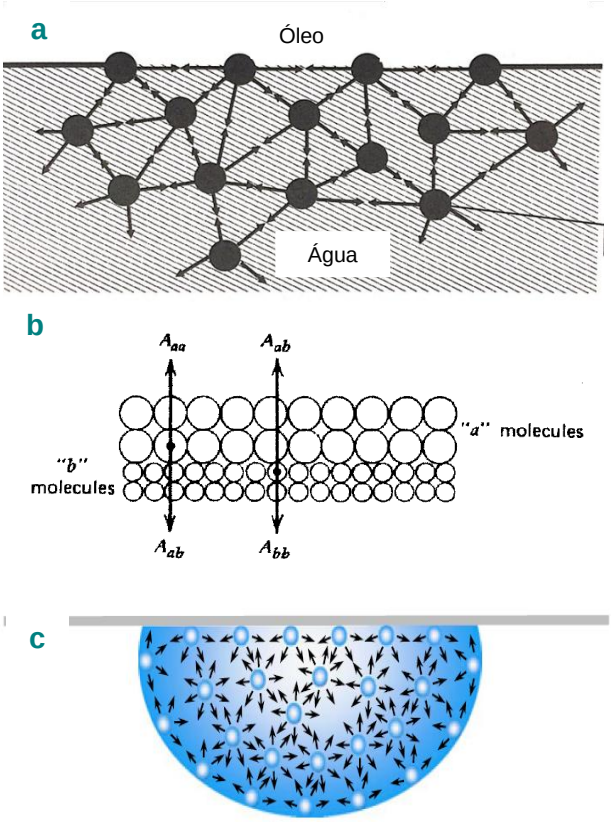

Figura 20 – Esquemas do desbalanço de forças intermoleculares e criação de interface entre fluidos. Fontes: Adaptado de [28], [29] e [30].

Essa diferença de energia potencial entre as moléculas da interface e as moléculas do corpo, por unidade de área da interface, é denominada tensão interfacial (no caso de dois líquidos) ou superficial (no caso de um líquido e um gás).

Os excessos de energias potenciais são dados pelas Equações 73a e 73b [29]:

$$\Delta E_A = A_{aa} - A_{ab} \ , \tag{73a}$$

onde ΔE_A é o excesso de energia da moléculas de A na interface em relação às do corpo e A_{aa}, A_{ab} são, respectivamente, a energia de interação entre as moléculas de A na interface com outras de A no corpo e a energia de interação entre moléculas de A e B na interface. Analogamente, para o excesso de energia das moléculas de B, temos

$$\Delta E_B = A_{bb} - A_{ab} \ . \tag{73b}$$

Então, o excesso de energia potencial de todas as moléculas na interface em relação a todas as moléculas do corpo, tanto de A quanto de B, é dado pela Equação 74.

$$\Delta E_{tot} = \Delta E_A + \Delta E_B = A_{aa} + A_{bb} - 2A_{ab} \tag{74}$$

Dividindo ΔE_{tot} por unidade de área da interface, obtemos a Equação 75.

$$\gamma = \frac{\Delta E_{tot}}{Área} = \frac{A_{aa} + A_{bb} - 2A_{ab}}{Área} = \gamma_a + \gamma_b - 2\gamma_{ab} \tag{75}$$

onde γ é a tensão interfacial e γ_a, γ_b e γ_{ab} são, respectivamente, a energia de interação “A com A” por unidade de área, a energia de interação “B com B” por unidade de área e a energia de interação “A com B” na interface por unidade de área.

Se um dos fluidos é gasoso, por exemplo o fluido B, as moléculas do mesmo encontram-se afastadas o suficiente para diminuir consideravelmente as interações intermoleculares, o que nos permite fazer as aproximações dadas pela Eq. 76.

$$\gamma_b = 0 \; e \; \gamma_{ab} = 0 \, . \tag{76}$$

Logo,

$$\gamma = \gamma_a \tag{77}$$

Ou seja, a tensão dependerá somente das forças de interação entre as moléculas de A e é então chamada de superficial.

Quando ambos os fluidos são não-gasosos, a tensão interfacial será dada pela Eq. 77. Se as polaridades[14] das moléculas de A e B forem semelhantes, a energia de interação entre elas (γ_{ab}) será grande. Observe que isto implica numa redução da tensão interfacial, conforme esperado: fluidos semelhantes não tendem a se segregar e formar interfaces bem definidas.

Uma mistura de água com álcool (altamente polares), por exemplo, fornecerá uma tensão interfacial baixa se comparada à mistura de água com óleo, ou seja, polar com apolar. A tensão interfacial água-óleo têm valores médios entre 20 e 40 mN/m, enquanto que a tensão água-álcool varia entre 5 e 10 mN/m [27].

A polaridade das moléculas, medida em Debyes, é dada por momentos de dipolo diferentes de zero e está relacionada, entre diversos outros fatores, à geometria molecular, polaridade de ligações e eletronegatividade dos átomos que a compõem (vide Fig. 21).

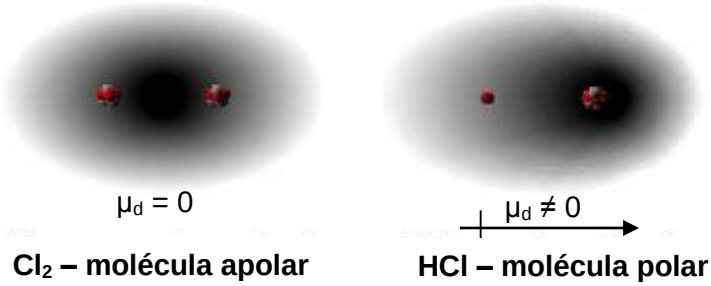

Figura 21 - Exemplo de uma molécula apolar, o gás cloro (Cl_2), e polar, o ácido clorídrico (HCl). Entre o átomo de hidrogênio e átomo de cloro há uma grande diferença de eletronegatividade, o que gera um momento de dipolo (μ_d) na molécula de HCl. Fonte: Adaptado de [31].

[14] A polaridade de moléculas pode ser definida como a presença de um momento de dipolo permanente diferente de zero.

De acordo com a polaridade das moléculas presentes em um sistema, haverá diferentes interações intermoleculares possíveis. Portanto, o que está por trás da “semelhança” entre fluidos na realidade são as forças intermoleculares e a intensidade das mesmas, não a polaridade em si. No exemplo dado, moléculas de água e álcool, por serem polares, podem realizar ligações de hidrogênio (as mais fortes depois das forças íon-dipolo), forças de dipolo-dipolo (relativamente fortes) e forças de dispersão de London (muito fracas) entre si. Portanto, quando em contato, interagem fortemente (são “semelhantes”). Não é o caso das moléculas de água e óleo, que podem realizar somente forças de dispersão de London (as únicas possíveis entre moléculas apolares e polares) [32].

3.1. Molhabilidade

Molhabilidade é a adesão relativa de dois fluidos em uma superfície sólida. Considerando dois fluidos imiscíveis em meios porosos, como é o caso de reservatórios de hidrocarbonetos, a molhabilidade é uma medida da tendência de um dos fluidos “molhar” (espalhar, aderir) preferencialmente as superfícies intersticiais do meio poroso na presença de outros fluidos [1].

As superfícies dos poros das rochas têm uma grande variedade de minerais expostos que podem ter afinidades por água, hidrocarbonetos ou por constituintes dissolvidos e suspensos nos fluidos (como os asfaltenos em óleos pesados, por exemplo). Inclusive este último caso pode ser responsável por mudanças na molhabilidade de reservatórios após processos de recuperação avançada (EOR).

Há fundamentalmente quatro estados gerais de molhabilidade em um reservatório ou corpo rochoso [1]:

- Molhável à água: ocorre quando a rocha em contato com os fluidos possui mais de 50% de sua superfície contatada pela água.
- Molhabilidade fracionada: ocorre quando há uma distribuição aleatória de molhabilidades ao longo de um corpo rochoso. Dependendo da distribuição de minerais expostos aos fluidos de formação, pode haver porções molháveis à água e outras molháveis ao óleo.
- Molhabilidade mista: ocorre quando os poros menores de uma rocha são molháveis à água e estão preenchidos com água, enquanto que os poros maiores são molháveis ao óleo e estão preenchidos com óleo. O óleo

não penetra nos poros menores devido à grande pressão capilar[15] necessária para o deslocamento da água.

- Molhável ao óleo: ocorre quanto a rocha em contato com os fluidos possui mais de 50% de sua superfície contatada pelo óleo.

Há ainda um estado complementar de molhabilidade intitulado “molhabilidade neutra”, termo que engloba as molhabilidades fracionada e mista. Em termos gerais, é utilizado quando metade na superfície rochosa é molhável à água e a outra metade molhável ao óleo, porém sem distinção entre tipos supracitados.

Em um reservatório trifásico (água, óleo e gás) molhável à água, por exemplo, observa-se que o fluido/fase molhante adere às paredes dos poros, isto é, permanece junto aos grãos de sólidos que compõe a rocha. Essa adesão, juntamente aos fenômenos capilares, origina a saturação irredutível de água, ou seja, a fração mínima de água restante após um processo completo de drenagem em um reservatório. As fases não-molhantes, neste caso o óleo e o gás, permanecem nas porções centrais dos poros, o que facilita o escoamento desses fluidos. A Figura 22 ilustra este fato.

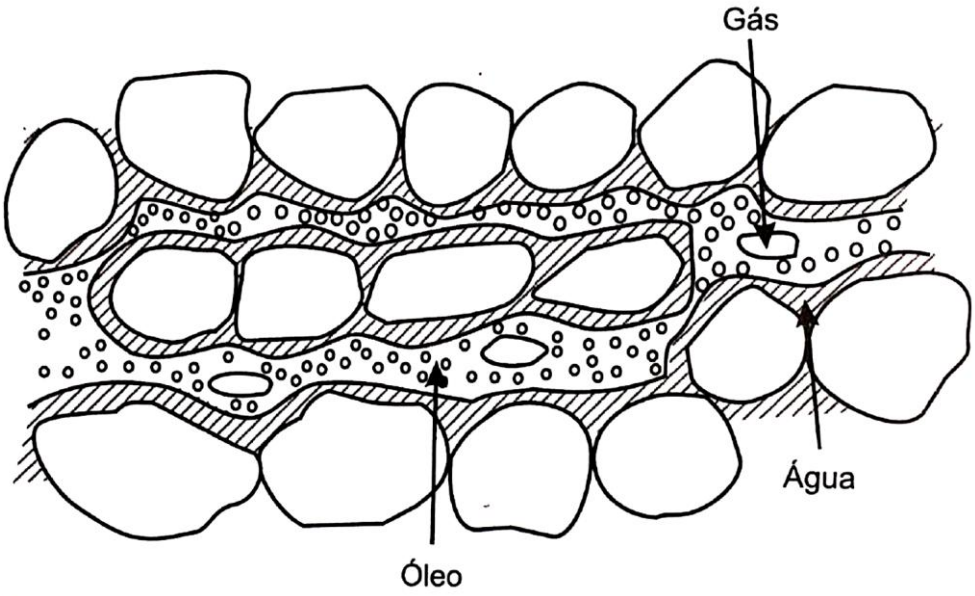

Figura 22 - Esquema de um reservatório molhável à água. Fonte: [28].

[15] A pressão capilar é definida como a pressão da fase não-molhante subtraída da pressão da fase molhante. É um fator determinante na saturação de fluidos de um reservatório e nos processos de embebição (aumento de saturação da fase molhante) e drenagem (aumento de saturação da fase não-molhante) que ocorrem durante as fases de migração de fluidos da rocha geradora à rocha reservatório de um sistema petrolífero.

Empiricamente, sabe-se que sedimentos siliciclásticos e minerais como o quartzo e o feldspato têm mais afinidade pela água, enquanto que sedimentos carbonáticos e minerais como calcita, dolomita, aragonita e siderita têm mais afinidade pelo óleo. Portanto, espera-se que a molhabilidade de arenitos, rochas-reservatório siliciclásticas, varie de neutra a fortemente molhável à água. Já quanto à molhabilidade de carbonatos (rochas-reservatório carbonáticas) espera-se que varie de neutra a molhável ao óleo [1]. Todavia, esses comportamentos não são uma regra: há, por exemplo, arenitos molháveis ao óleo por causa de interações de seus minerais com compostos polares de óleos de formação.

Segundo DONALDSON & ALAM [1], a molhabilidade é o principal parâmetro que afeta: a distribuição microscópica da água e do óleo em poros de rochas; as características das curvas de pressão capilar e, consequentemente, a permeabilidade relativa[16] dos fluidos; as propriedades do escoamento de óleo sendo varrido[17] por água; e a quantidade e distribuição da saturação residual[18] de óleo.

3.2. Ângulo de Contato

Considere uma gota de óleo em contato com uma amostra pequena de rocha e imersa em um ambiente de salmoura (água salina). A gota poderá adotar uma configuração entre espalhar-se completamente na superfície do sólido (superfície infinitamente molhável ao óleo, com ângulo de contato igual a 180º) ou repousar em formato de gota completamente redonda (superfície infinitamente molhável à água, com ângulo de contato igual a 0º). Entre esses dois extremos, o formato da gota exibirá um ângulo de contato mensurável, geralmente medido em relação à fase mais densa [1], neste caso, a salmoura.

[16] A permeabilidade relativa é definida como o quociente entre a permeabilidade efetiva a um fluido (permeabilidade de uma rocha a um determinado fluido quando saturada por mais de um fluido) e a permeabilidade absoluta.
[17] O termo “varrido” é muito comum na indústria petrolífera e indica o deslocamento de um fluido no reservatório por outro fluido, seja ele injetado ou já presente no reservatório.
[18] Fração de óleo no reservatório não recuperada após produção. Métodos de recuperação melhorada (IOR) e avançada (EOR) de petróleo têm como principal objetivo a sua redução.

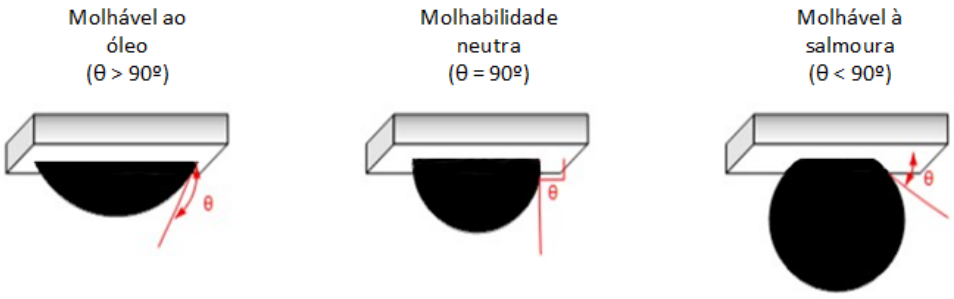

Figura 23 – Esquema ilustrando ângulos de contato (θ) e molhabilidades correspondentes. Considere o óleo imerso em um ambiente de salmoura.

A tensão interfacial óleo-salmoura (γ_{os}) age tangencialmente à superfície da gota. No equilíbrio de forças na interface óleo-salmoura-rocha, temos a situação representada pela Figura 24, onde γ_{or} e γ_{sr} são, respectivamente, a tensão interfacial óleo-rocha e salmoura-rocha.

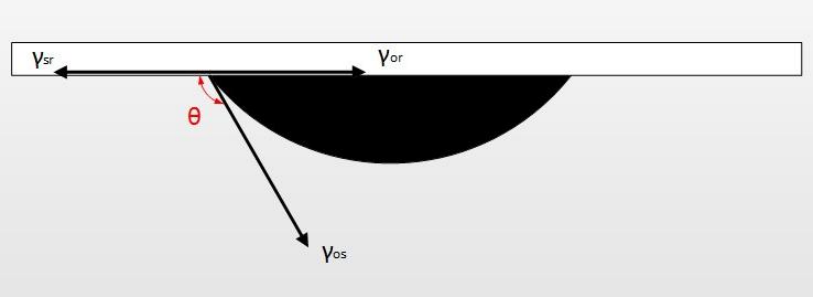

Figura 24 - Ângulo de contato e equilíbrio de forças na interface entre os fluidos. A cor preta simboliza o óleo e a cor cinza, a salmoura.

A histórica equação de Thomas Young (Eq. 78), exímio cientista britânico do século XIX, descreve o sistema da Figura 24 [28]:

$$\gamma_{ads} = \gamma_{or} - \gamma_{sr} = \gamma_{os} \cos\theta \,, \tag{78}$$

onde θ é o ângulo de contato de Young e γ_{ads} é a tensão de adesão, *i.e.* a diferença das tensões ao longo do sólido. Ela determina qual fluido molha preferencialmente o sólido [28].

Aplicando simples passos algébricos e a função inversa do cosseno na Eq. 78, obtém-se uma relação explícita para o ângulo de contato:

$$\theta = \cos^{-1}\left(\frac{\gamma_{or} - \gamma_{sr}}{\gamma_{os}}\right) \quad (79)$$

Com relação à determinação experimental de ângulos de contato, há algumas técnicas bem desenvolvidas. Segundo BRACCO & HOLST [33], a maioria das técnicas pode ser subdividida em dois grandes grupos: métodos ópticos (diretos) e métodos de força (indiretos). A medição pelo *Drop Shape Analyzer* (DSA) é um exemplo de método óptico, assim como a medição por telescópio-goniômetro (Figura 25a) e pelo método de placa oblíqua (Figura 25b). Como exemplos de métodos indiretos, pode-se citar: o método de Wilhelmy (Figura 25c), baseado na variação da força peso de uma amostra quando em contato com fluidos devido a forças de empuxo e forças resultantes da molhabilidade; e a medição por ascensão capilar (Figura 25d).

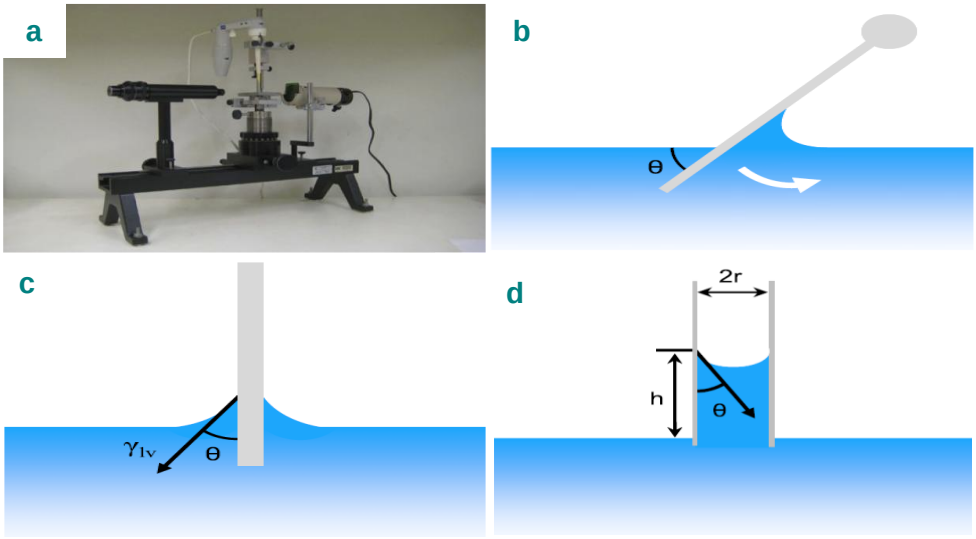

Figura 25 - Métodos diretos (a e b) e indiretos (c e d) de medição experimental de ângulos de contato. Fonte: [30].

4. Medição de Ângulos de Contato: Metodologia e Procedimentos

4.1. O *Drop Shape Analyzer*

O *Drop Shape Analyzer* (DSA), conhecido no Brasil como tensiômetro de gota pendente[19] ou gota séssil[20], é um equipamento de tecnologia óptica para medição de tensões interfaciais e ângulos de contato. Acoplado a um *software* de captura e processamento de imagens, permite realizar análises de forma objetiva e automatizada (Figura 26).

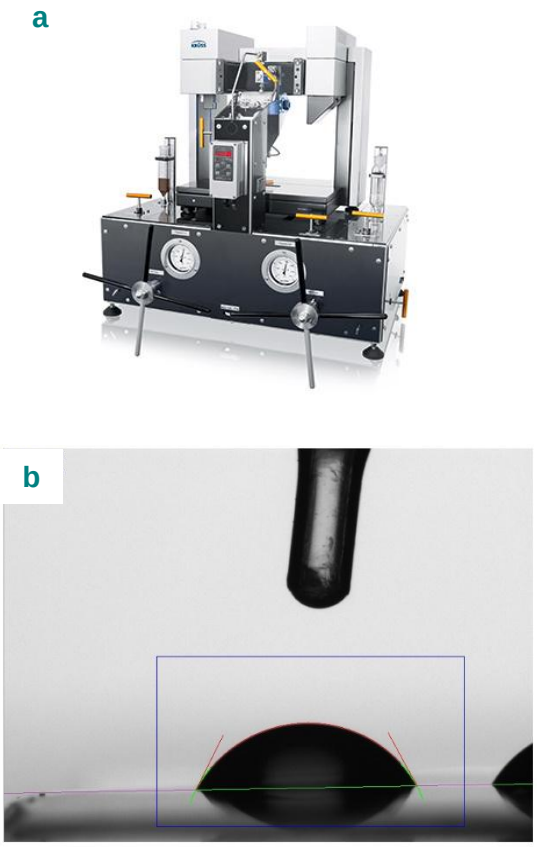

Figura 26 - *Drop Shape Analyzer DSA100HP*® da Krüss (acima) e um exemplo de medição de gota sessil utilizando o *software DSA4*® e o Método da Tangente (abaixo). Fonte: Adaptado de [30] e [34].

O conjunto do equipamento é composto basicamente de um sistema óptico, uma câmera de alta resolução, uma câmara de observação resistente à alta pressão e

[19] Quando o fluido estudado é menos denso que o fluido em que está imerso.
[20] Quando o fluido estudado é mais denso que o fluido em que está imerso.

temperatura, uma fonte luminosa, sensores de temperatura e pressão, válvulas, linhas de 1/4 a 1/8 de polegada e controlador de temperatura. A Figura 27 ilustra esses componentes.

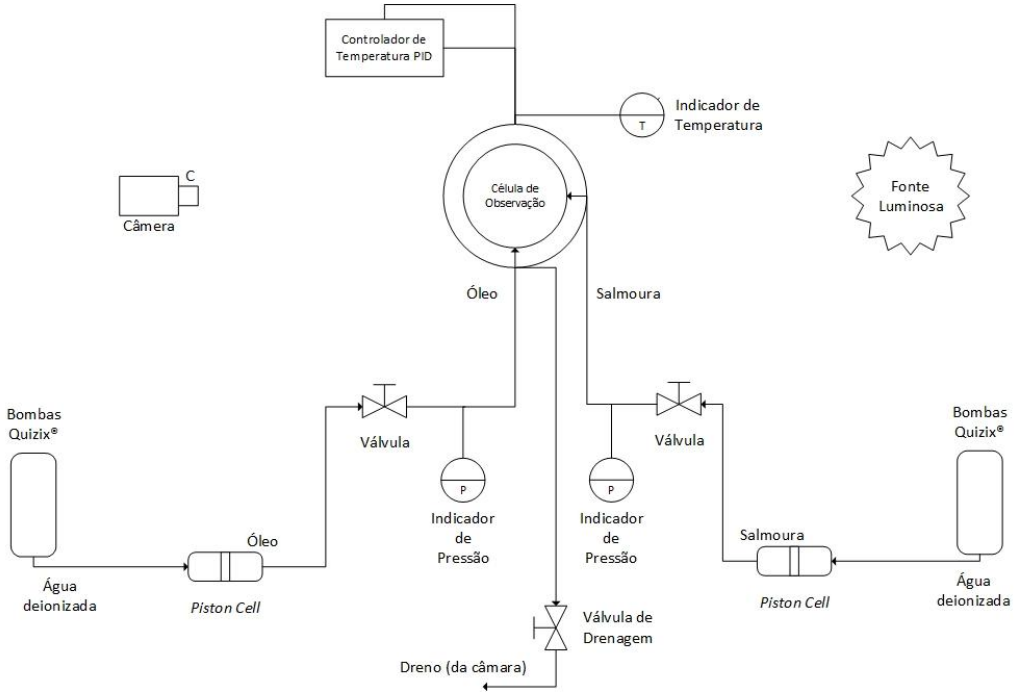

Figura 27 - Esquema dos componentes do DSA, acrescido de equipamentos auxiliares (Bombas Quizix® e recipientes de fluidos).

Em todos os experimentos realizados, o *DSA100HP*® e o *software DSA4*® da empresa alemã Krüss foram utilizados [34]. O equipamento estava acoplado às bombas *Quizix® Q5000* da empresa norte-americana Chandler Engineering e células acumuladoras de fluidos com pistão interno que suportam até 15.000 psi (*piston cells*).

4.2. Metodologia Experimental

Para a medição estática de ângulos de contato (ACs), que é o caso deste trabalho, uma gota de óleo é injetada na linha que conecta uma célula acumuladora até a câmara de observação. A gota, que se desprende da linha e adere a uma amostra de sólido (rocha/mineral) imersa em salmoura, é iluminada de um lado e a câmera é posta do lado oposto, capturando imagens da gota. A Figura 28 ilustra o processo.

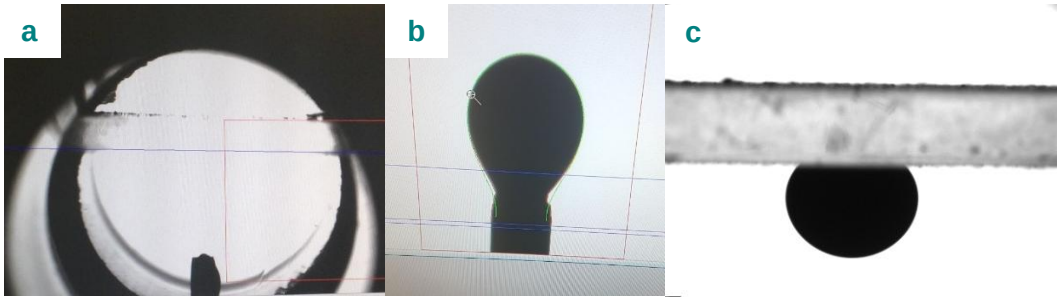

Figura 28- A imagem da esquerda ilustra a visão no *software DSA4*® da câmara de observação. É possível observar a linha de injeção no canto inferior da imagem. A imagem central ilustra a injeção de uma gota de óleo. A imagem da direira mostra um quartzo e o óleo em contato com o mineral, imersos em salmoura e prontos para a medição.

Utilizando o DSA, o procedimento mais comum de cálculo do ângulo de contato é o Método da Tangente: o *software* de processamento de imagens reconhece a forma da gota a partir de um algoritmo baseado em escalas de cinza [34] e, então, o usuário ajusta uma linha-base (*baseline*) que passe pelo ponto triplo (contato de todas as fases, fluidos e sólido) em ambos os lados da gota para que se trace uma reta tangente a ela. A Figura 29 ilustra o procedimento de ajuste da baseline.

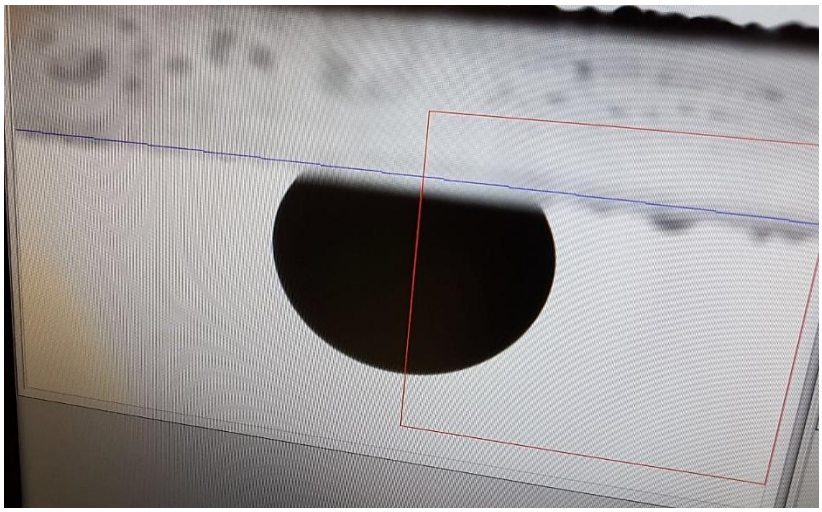

Figura 29 - Ajuste da *baseline* (linha azul) no *software* DSA4. O quadrado vermelho representa a área focal: no caso, buscava-se medir o ângulo direito da gota pendente de óleo imersa em no ambiente de salmoura.

A incerteza do Método da Tangente geralmente está entre ±1º e ±2º, porém dependendo da molhabilidade do sólido esses valores podem variar significativamente [35].

É importante ressaltar que, apesar de as medições serem estáticas e de as gotas injetadas terem volumes aproximadamente constantes, os ângulos de contatos não permanecem constantes sempre ao longo dos testes. Fenômenos na superfície de contato entre os fluidos ou entre fluidos e sólido, *e.g.* volatilidade de frações mais leves do fluido injetado ou ainda reações químicas entre o sólido e o fluido, podem fazer com que os ACs variem com o tempo (vide Figura 30).

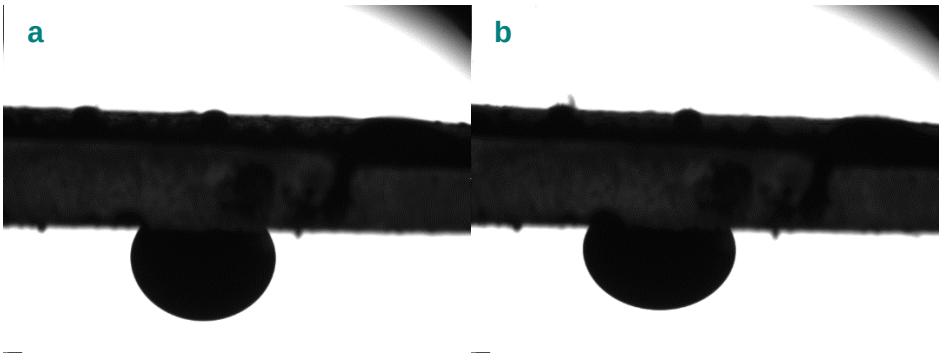

Figura 30 - A imagem da esquerda representa o início de medição de ângulos de contato de um substrato de dolomita com óleo, imersos em salmoura. A imagem da direita ilustra o mesmo sistema ao final das medições. Observe a diferença de formato da gota e de ângulo de contato. O ângulo de contato médio no início da medição era de 59,4º, enquanto que ao final era de 63,7º.

Todos os experimentos realizados para a criação deste trabalho tiveram a duração de um dia (vinte e quatro horas). Devido às incertezas do Método da Tangente e também ao período de variação do AC, somente as medidas estáveis, isto é, as vinte últimas medidas em que o AC não variava mais que dois graus para cima ou para baixo ($-2^o \leq \Delta \leq +2^o$), foram levadas em consideração nas análises. Além disso, os resultados de ângulo de contato apresentados na seção 5 são os valores médios entre os valores dos ângulos esquerdo e direito de cada gota.

O óleo utilizado nas medições era proveniente de um importante reservatório do pré-sal brasileiro que não pode ser nomeado por motivos de confidencialidade. A salmoura deste mesmo campo, cuja composição e concentração de íons eram conhecidas, foi reproduzida em laboratório com altíssima fidelidade utilizando solutos de alta pureza da empresa norte-americana Sigma-Aldrich. As Tabelas 12 e 13 contém as principais informações dos fluidos de formação [36].

Tabela 12 – Características principais da salmoura de formação reproduzida em laboratório. Fonte: [36].

Massa Específica [kg/cm³]*	Composição		
	Íon	Concentração [pmm]	TDS** [ppm]
1.172,30	Na^{+}	57.580,00	226.694,10
	Ca^{2+}	24.250,00	
	Mg^{2+}	2.120,00	
	K^{+}	1.200,00	
	Ba^{2+}	24,00	
	Sr^{2+}	1.260,00	
	SO_4^{2-}	1.260,00	
	Cl^{-}	139.954,06	

* 60 ºC, 1.000 psi.
** *Total Dissolved Solids* (Total de Sólidos Dissolvidos).

Tabela 13 - Características principais do óleo morto do reservatório do pré-sal brasileiro. Fonte: [36].

Massa Específica [kg/cm³]*	Composição	
	SARA**	Concentração [%]
849,20	Saturados	58,82
	Aromáticos	21,38
	Resinas	16,07
	Asfaltenos	3,73

*60 ºC, 1.000 psi.
**Saturados, Aromáticos, Resinas e Asfaltenos

Durante aproximadamente quatro meses, foram realizados vinte e quatro experimentos variando os substratos sólidos (minerais) e os solventes utilizados para limpeza (vide seção 4.2 para maiores detalhes). Os minerais, apesar de oriundos de afloramentos rochosos localizados nos Estados Unidos da América, são minerais representativos de dois tipos de reservatórios do pré-sal: areníticos e carbonáticos. A Tabela 14 resume todos os testes feitos.

Tabela 14 - Experimentos de ângulo de contato realizados. O número do experimento não reflete ordem cronológica de realização: os experimentos foram realizados em ordem aleatória.

Mineral	Experimento	Envelhecimento na Salmoura (dias)	Envelhecimento no Óleo (dias)	Solvente
Quartzo	1	0	0	Nenhum
	2	1	3	Salmoura
	3	1	3	N-Hexano
	4	1	3	Ciclohexano
	5	1	3	Hexadecano
	6	1	3	Tolueno
	7	1	15	Ciclohexano
	8	1	30	Ciclohexano
Calcita	9	0	0	Nenhum
	10	1	3	Salmoura
	11	1	3	N-Hexano
	12	1	3	Ciclohexano
	13	1	3	Hexadecano
	14	1	3	Tolueno
	15	1	15	Ciclohexano
	16	1	30	Ciclohexano
Dolomita	17	0	0	Nenhum
	18	1	3	Salmoura
	19	1	3	N-Hexano
	20	1	3	Ciclohexano
	21	1	3	Hexadecano
	22	1	3	Tolueno
	23	1	15	Ciclohexano
	24	1	30	Ciclohexano

Portanto, os vinte e quatro experimentos realizados e as vinte medidas tomadas em cada um totalizaram 480 medições de AC (com amostras de vinte medições para cada subpopulação/experimento), um número suficiente para realizar análises estatísticas robustas e confiáveis (vide Figura 31).

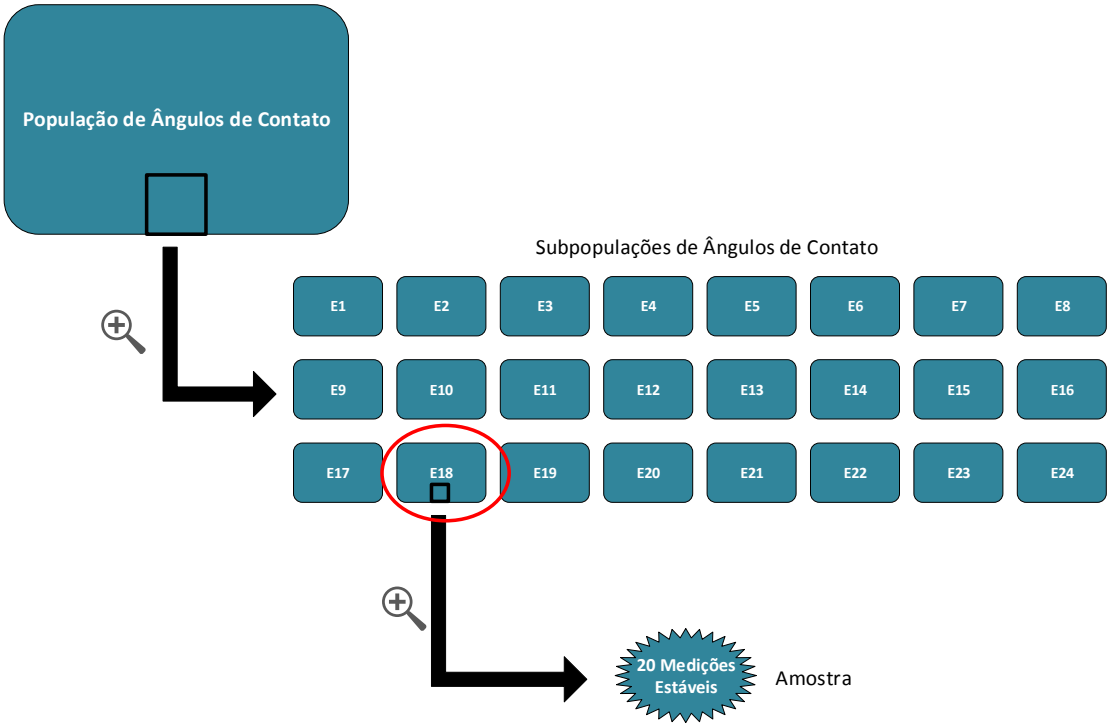

Figura 31 - Esquema ilustrativo dos conceitos estatísticos envolvidos no trabalho. A partir de uma população de ângulos de contato (todos os ângulos de contatos possíveis), selecionam-se subpopulações de ACs, ou seja, subconjuntos dos valores possíveis (subpopulações). Estes subconjuntos são os vinte e quatro experimentos (simbolizados por "E") realizados, cada um com suas características (minerais, solventes e envelhecimento). Dentre cada subpopulação, foram selecionadas vinte medições (amostra). A figura ilustra o processo para o experimento 18 (E18). Vinte medições em cada um dos vinte e quatro experimentos totalizaram 480 medições.

Tanto para os minerais quanto para os solventes, foram escolhidas características representativas que traduzissem as variáveis classificatórias em variáveis numéricas, tornando possível a realização das análises estatísticas. No caso dos minerais, o Ponto de Zero Carga (pzc), ou seja, o pH em que um sólido submerso em um eletrólito exibe carga elétrica líquida nula, foi utilizado. Já no caso dos solventes, a Constante Dielétrica (ε), que pode ser relacionada à polaridade, foi utilizada. Foram selecionados valores médios da literatura (Tabelas 15 e 16). Não houve medição em laboratório das propriedades supracitadas.

Tabela 15 - Substratos rochosos e seus respectivos pontos de zero carga.

Mineral	pzc	Fonte
Quartzo	3,0 (25 °C)	Kosmulski, 2002 [37]
Calcita	9,5 (20 °C)	Al Shalabi e Sepehrnoori, 2017 [38]
Dolomita	7,4 (20 °C)	Al Shalabi e Sepehrnoori, 2017 [38]

Tabela 16 - Solventes e suas respectivas constantes dielétricas.

Solventes	ε	Fonte
Nenhum (Ar)	≅ 1,000 (25 °C)	Hector e Schultz, 1936 [39]
N-hexano	1,890 (20 °C)	Lide, 2005 [40]
Ciclohexano	2,015 (25 °C)	Maryott e Smith, 1951 [41]
N-Hexadecano	2,090 (25 °C)	Aralaguppi *et al.*, 1991 [42]
Tolueno	2,379 (25 °C)	Maryott e Smith, 1951 [41]
Salmoura	78,540* (25 °C)	Maryott e Smith, 1951 [41]

*Constante dielétrica da água.

Note que a constante dielétrica da água foi utilizada, visto que um valor para a salmoura do campo do pré-sal não foi determinado em laboratório. Esta escolha não afeta em nenhum ponto as análises estatísticas da seção 5, pois as técnicas se baseiam em variações entre os dados e na determinação de padrões e correlações. Observe que, tanto para a água quanto para a salmoura, o valor de ε seria o maior de todos os solventes, pois são muito polares. O que diferiria seria que o valor para a salmoura seria ainda maior que o da água, pois os íons em solução daquela diminuem a resistência à criação de um campo elétrico.

As condições termodinâmicas, fixas para todos testes, estão descritas na Tabela 17.

Tabela 17 - Variáveis termodinâmicas dos testes.

Variável Termodinâmica	Valor
Pressão [psi]	1.000,00
Temperatura [°C]	60,00

4.3. Preparação de Amostras

Antes de as amostras de minerais, as quais possuem dimensões aproximadas de 25 mm x 25 mm x 3 mm e são conhecidas como *rock slabs*, serem colocadas na câmara de observação, as mesmas precisam ser preparadas.

A preparação de amostras para os testes de AC consistiu basicamente em duas etapas: envelhecimento e limpeza. O envelhecimento (*aging*) é o tempo em que a amostra de sólido, imersa em um becker com salmoura e depois em outro com óleo, era colocada em uma estufa com temperatura fixa (no caso, 60 ^{o}C). Para a salmoura, esse tempo variou entre zero e um dia. Para o óleo, entre zero, um, três, quinze e trinta dias. O objetivo é simular os longos processos de embebição e drenagem de um reservatório de hidrocarbonetos, o qual presume-se que esteja saturado com salmoura até que o óleo advindo da rocha geradora migre até o mesmo. A Figura 32 ilustra o procedimento.

Figura 32 - O processo de envelhecimento de amostras.

Já a etapa de limpeza consiste em retirar o excesso de óleo da superfície da amostra após o envelhecimento em óleo na estufa. Sem a limpeza, a gota injetada no DSA espalharia na superfície do mineral e não seria possível medir o AC. Dentre os diversos modos de limpeza existentes, dois são mais comuns: centrífuga e uso de solventes. O último caso foi escolhido para o estudo deste projeto (a Tabela 16 da seção 4.2 descreve quais foram utilizados). Para tanto, despejavam-se com uma pipeta três gotas do solvente em cada lado do *rock slab*, espalhando-as na superfície o máximo possível. Em seguida, a amostra recebia um banho de salmoura, a mesma a ser utilizada no teste. A Figura 33 ilustra o procedimento.

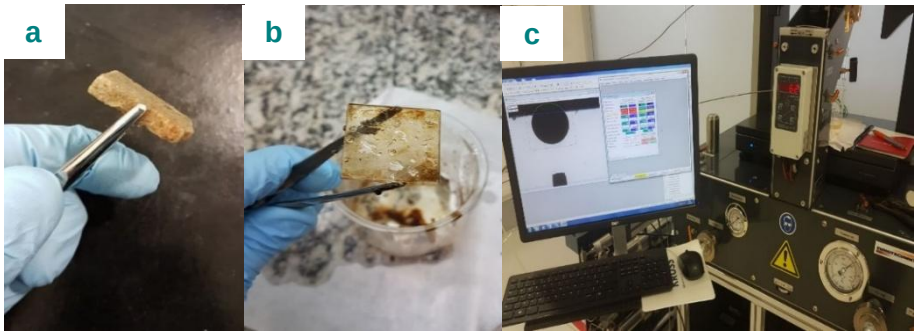

Figura 33 – Um *rock slab* cru, sem envelhecimento (esquerda). Uma amostra após o envelhecimento em óleo, limpeza com solvente e banho de salmoura (centro). O *DSA100HP*® e o início de um experimento (direita).

5. Resultados Experimentais e Análise Estatística

Os resultados experimentais médios obtidos nos vinte e quatro testes realizados, seguindo o critério explicitado na seção 4.2, estão compilados na Tabela 19 da próxima página (vide Anexo III para todas as medições). Observe que em um dos experimentos, o de número 3 na Tabela 19, a gota espalhou-se na superfície do *rock slab* e não foi possível obter uma medição de ângulo de contato (a Fig. 34 ilustra o ocorrido). Portanto, o número total da população após o incidente é, na realidade, 460.

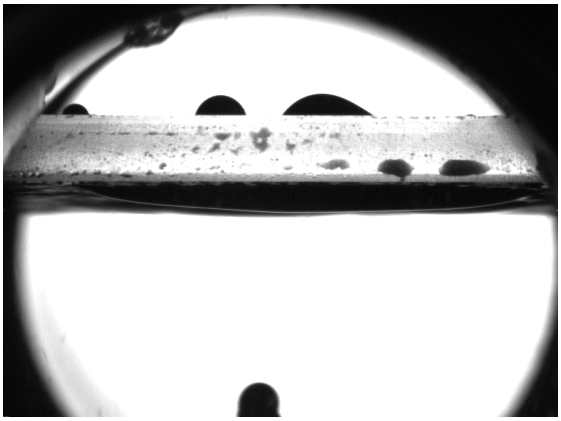

Figura 34 - Gota de óleo espalhada em uma amostra de quartzo limpa com n-hexano.

Todas as análises estatísticas foram realizadas com o auxílio do Statsoft *Statistica 7.0*®, um *software* focado em Estatística que realiza desde análises simples até extremamente complexas e em grandes escalas.

As medidas de posição, dispersão, assimetria e curtose para a população de ângulos de contato são apresentadas na Tabela 18.

Tabela 18 - Medidas de posição, dispersão, assimetria e curtose da população de ângulos de contato.

Média [º]	90,7902
Mediana [º]	88,9500
Moda	Múltipla
Valor Mínimo [º]	49,0000
Valor Máximo [º]	163,2000
Variância [º^2]	949,2416
Desvio-Padrão [º]	30,8098
Coef. de Assimetria	0,6762
Coef. De Curtose	-0,3446

Tabela 19 - Resultados experimentais de ângulo de contato.

Mineral	Experimento	Envelhecimento na Salmoura (dias)	Envelhecimento no Óleo (dias)	Solvente	Resultado de Ângulo de Contato* (º)	Coeficiente de Variação**
Quartzo	1	0	0	Nenhum	81,7	0,21%
	2	1	3	Salmoura	92,8	0,08%
	3	1	3	N-Hexano	-	-
	4	1	3	Ciclohexano	64,4	0,22%
	5	1	3	Hexadecano	107,0	0,13%
	6	1	3	Tolueno	59,5	0,15%
	7	1	15	Ciclohexano	143,7	0,39%
	8	1	30	Ciclohexano	88,9	0,21%
Calcita	9	0	0	Nenhum	51,9	0,41%
	10	1	3	Salmoura	77,2	0,38%
	11	1	3	N-Hexano	71,5	0,17%
	12	1	3	Ciclohexano	61,4	1,39%
	13	1	3	Hexadecano	92,2	0,14%
	14	1	3	Tolueno	100,1	0,10%
	15	1	15	Ciclohexano	162,8	0,19%
	16	1	30	Ciclohexano	113,6	0,68%
Dolomita	17	0	0	Nenhum	50,3	1,06%
	18	1	3	Salmoura	82,6	0,32%
	19	1	3	N-Hexano	129,0	0,41%
	20	1	3	Ciclohexano	63,9	0,35%
	21	1	3	Hexadecano	102,7	0,69%
	22	1	3	Tolueno	52,7	0,17%
	23	1	15	Ciclohexano	95,6	0,64%
	24	1	30	Ciclohexano	142,9	0,30%

* Média aritmética das últimas vinte medições seguindo o critério adotado.

** Calculado somente para as últimas vinte medições seguindo o critério adotado.

É possível observar na Tabela 18 que a distribuição de valores de ângulos de contato é: multimodal (tem mais de uma moda); aproximadamente simétrica (coeficiente de assimetria próximo a zero e média muito próxima à mediana); platicúrtica (coeficiente de curtose menor que zero); e de grande amplitude (114 graus).

Qualitativamente, é possível perceber alguns padrões na Tabela 19: o quartzo apresenta menores valores de ângulo de contato se comparado aos carbonatos, indicando que aquele é molhável à salmoura; quando solventes saturados como o N-hexano e o Hexadecano foram utilizados para limpeza, os resultados indicam ângulos de contato maiores que 90º (molhabilidade ao óleo). Observando a composição do óleo na Tabela 13 da seção 4.2, mais da metade dos componentes do óleo são hidrocarbonetos saturados, o que explica a tendência observada, visto que o óleo interage fortemente com o solvente; muitos valores encontram-se na faixa de molhabilidade neutra (*i.e.* 80º-100º) para a maioria dos solventes.

A primeira análise quantitativa realizada foi a construção da matriz de correlações. Seus valores são os coeficientes de correlação de Pearson (CCP) para cada par *ij* de variáveis (vide seção 2.2.2 para mais detalhes). Utilizada como ferramenta de estudo preliminar, seu objetivo maior é confirmar padrões já esperados ou intuições construídas de acordo com o conhecimento dos princípios dos fenômenos envolvidos nos experimentos e/ou de acordo com informações da literatura. Por exemplo, esperava-se que quanto maior o tempo de envelhecimento no óleo, maior o ângulo de contato, ou seja, mais molhável ao óleo ($\theta > 90º$) seria o mineral. A Tabela 20 e a Figura 35 contêm a matriz de correlações obtida para os dados experimentais.

Tabela 20 - Matriz de correlações dos dados experimentais. Valores em negrito indicam significância estatística considerando $\alpha = 5\%$. Nomenclatura: T_{es} é o tempo de envelhecimento em salmoura, T_{eo} o tempo de envelhecimento em óleo e *AC* é o ângulo de contato.

	Mineral	**T_{es}**	**T_{eo}**	**Solvente**	**AC**
Mineral	1,00	0,02	-0,03	-0,02	0,00
T_{es}	0,02	1,00	**0,31**	**0,16**	**0,37**
T_{eo}	-0,03	**0,31**	1,00	**-0,19**	**0,56**
Solvente	-0,02	**0,16**	**-0,19**	1,00	-0,08
AC	0,00	**0,37**	**0,56**	-0,08	1,00

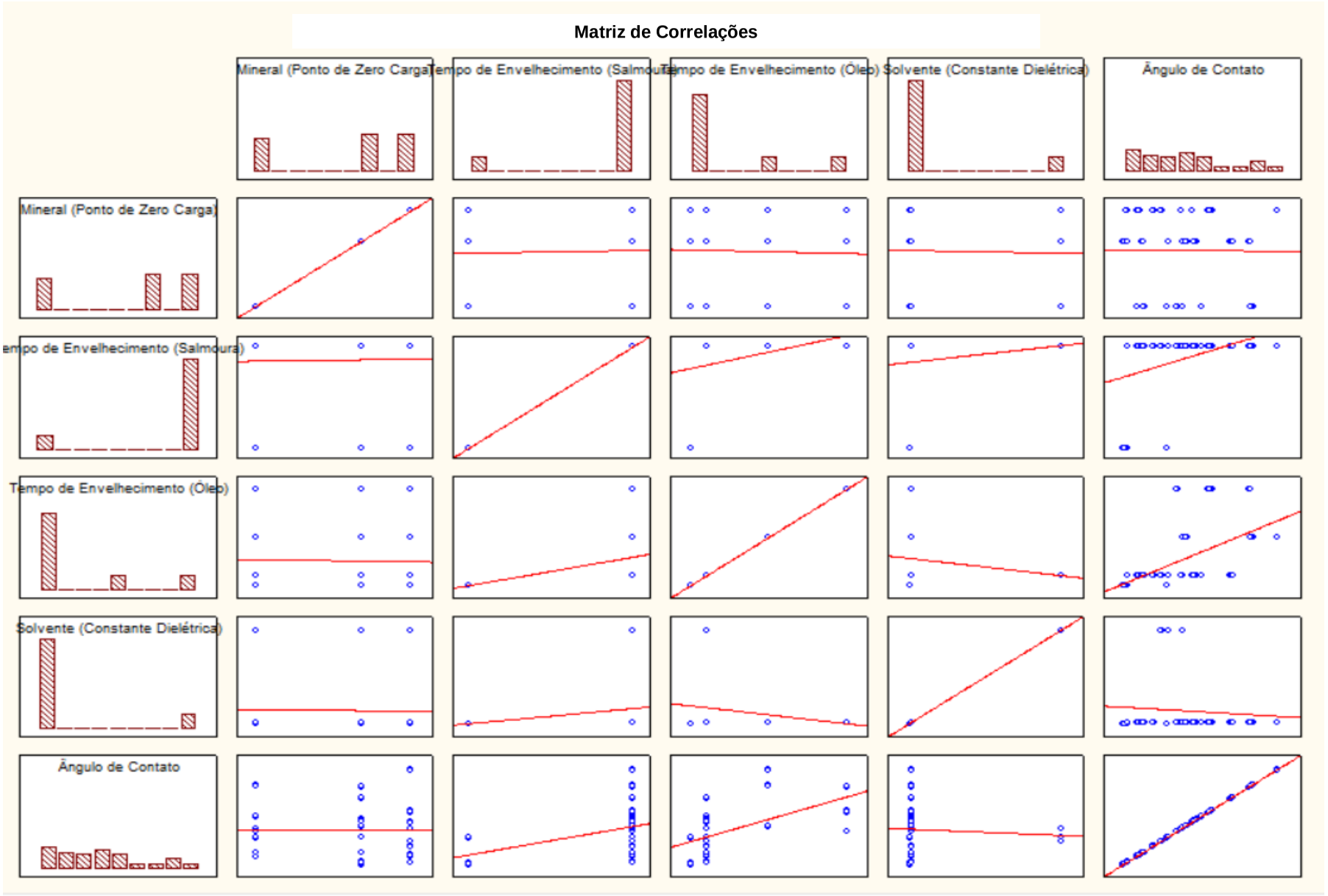

Figura 35 - Matriz de correlações no formato de gráficos de dispersão.

Na Tabela 20 e Figura 35, é claro observar que, segundo os CCP*s*, as variáveis com maior correlação com os resultados de ângulos de contatos são os tempos de envelhecimento em salmoura e em óleo, com destaque para o óleo (*CCP* = 0,56). Ambos são estatisticamente significantes com α = 5% e positivos, indicando que quanto maior o tempo de envelhecimento, maior o ângulo de contato.

Contraintuitivamente, a análise preliminar não encontrou correlação com os minerais e com os solventes. Contudo, os resultados para estas variáveis não foram estaticamente significantes e, portanto, não devem ser levados em consideração nesta primeira análise. Além disso, Fenômenos de Superfícies são altamente complexos e envolvem diversos fatores. Espera-se, portanto, que os resultados de AC apresentem não-linearidade.

A segunda análise estatística realizada foi a ANOVA via Regressão Linear Múltipla. Verificou-se previamente a normalidade da população com um NPP (Fig. 36). O gráfico demonstra que a população de ângulos de contato possui alguns *outliers* (nos cantos inferior esquerdo e superior direito) no ajuste à distribuição Normal, conforme era esperado: a Estatística Descritiva demonstrou que a distribuição de ângulos de contato é um pouco mais achatada que a Normal. Segundo ZAIONTZ [43], a premissa da normalidade pode ser relaxada se o tamanho das subpopulações é o mesmo (modelo balanceado) e se são suficientemente grandes (maiores que 10 no mínimo), como é o caso. Portanto, escolheu-se seguir com a análise.

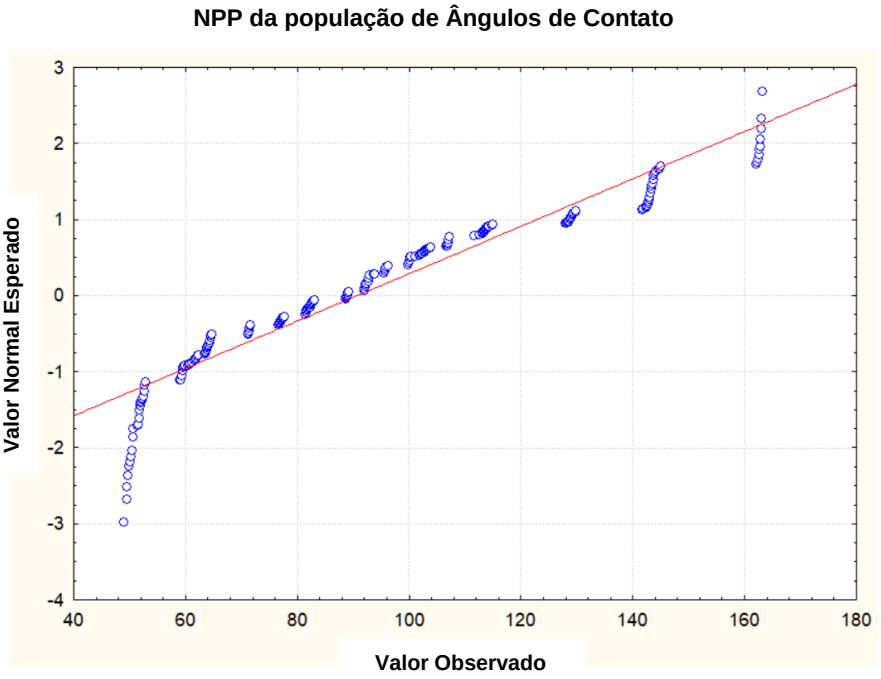

Figura 36 - NPP da população de 460 ângulos de contato determinados experimentalmente. A linha reta indica ajuste perfeito à distribuição Normal.

Quanto à condição de homocedasticidade (igual variância entre as subpopulações), a mais relevante para a ANOVA, observe a Tabela 20, que contém as variâncias de cada experimento, ou seja, de cada subpopulação de ângulos de contato. Na seção 2.3.4, vimos que as variâncias amostrais devem ser rigorosamente iguais. Contudo, há uma regra prática [43] que postula que se o quociente da maior variância dentre as subpopulaçãos pela menor das variâncias for menor que 4, a ANOVA ainda é robusta. Portanto, escolheu-se realizar a ANOVA utilizando somente as subpopulações que se adequavam à esta regra prática. Utilizou-se a variância do Experimento 2 como a base de comparação (variância mínima), visto que é a que fornece mais experimentos se adequando à regra prática (1, 2, 4 ao 6, 8 ao 11, 13 ao 15, 18, 20 e 22), totalizando 320 medições (vide Tabela 21). É importante ressaltar que essa escolha altera o Universo do estudo e que a estimação de parâmetros para as subpopulações excluídas constitui uma Indução Estatística.

Tabela 21 - Variância das subpopulações de ângulos de contato. Os campos em verde simbolizam os experimentos que respeitam a regra prática de variâncias com base na variância do Experimento 2.

Experimento/Subpopulação	Variância [$^{o^2}$]
1	0,0303
2	0,0060
3	-
4	0,0209
5	0,0206
6	0,0081
7	0,3149
8	0,0363
9	0,0453
10	0,0845
11	0,0145
12	0,7284
13	0,0163
14	0,0109
15	0,0955
16	0,6019
17	0,2860
18	0,0693
19	0,2740
20	0,0511
21	0,5009
22	0,0074
23	0,3746
24	0,1875

Finalmente, a última condição de subpopulações independentes é satisfeita, pois os experimentos foram realizados de forma aleatória. Então, a ANOVA via Regressão Linear Múltipla pôde ser efetuada. Antes da apresentação de qualquer resultado da regressão, é importante verificar se os resíduos obtidos seguem uma distribuição Normal, visto que a teoria assume que erros aleatórios não são sistemáticos e devem seguir essa distribuição de probabilidades. A Figura 37 demonstra a verificação.

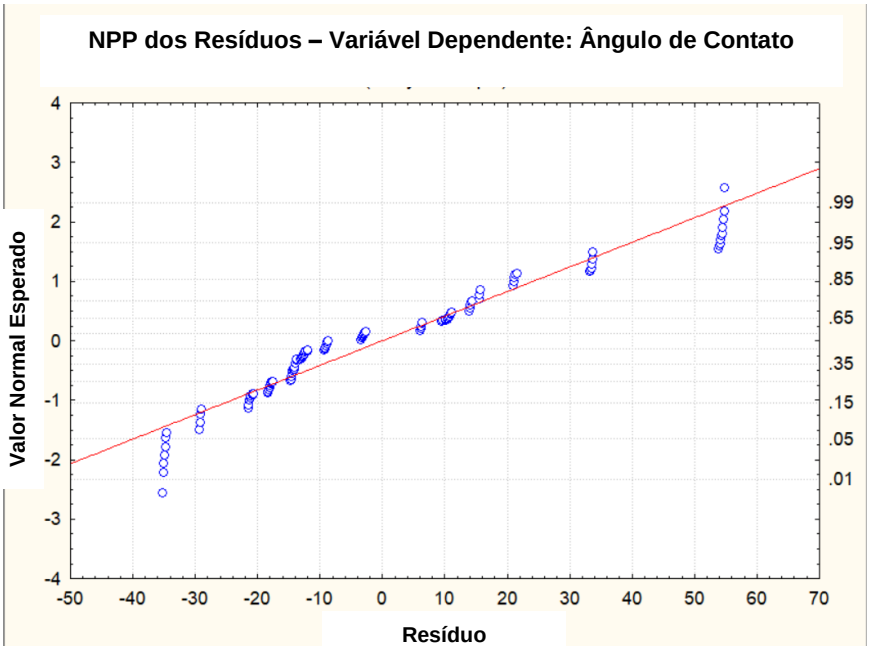

Figura 37 - Verificação da normalidade dos resíduos da Regressão Linear Múltipla. A linha reta é o ajuste perfeito à distribuição Normal. Observa-se apresença de alguns *outliers* nos extremos.

Através da Figura 37, é possível concluir que os resíduos apresentam alguns *outliers* à extrema esquerda e direita. Relaxou-se a premissa da normalidade seguindo a mesma regra prática dada por ZAIONTZ [43] e, então, prosseguiu-se com a análise. Realizando a ANOVA, foram obtidos os resultados da Tabela 22.

Tabela 22 - Tabela ANOVA via Regressão para os dados experimentais. α = 5%.

Fonte de Variação	SS	DOF	MS	F	valor-p
Mineral	**9627.1**	**1**	**9627.06**	**17.5665**	**0.000036**
T_{es}	1569.4	1	1569.35	2.8636	0.091593
T_{eo}	**83126.4**	**1**	**83126.38**	**151.6808**	**0.000000**
Solvente	629.5	1	629.52	1.1487	0.284646
Residual	172631.0	315	548.03		
Total	267583.3	319.0			

Novamente, é possível perceber o tempo de envelhecimento em óleo foi a variável mais influente, com o maior valor da estatística de teste *F*. A variável mineral também provou ser influente com significância estatística, atrás apenas do T_{eo}.

Quanto aos resultados da Regressão Linear Múltipla, foi encontrada, através do Método dos Mínimos Quadrados, a Eq. 80 como melhor ajuste. É importante destacar que ela é válida somente para dados experimentais que estejam nos mesmos intervalos das variáveis explicativas estudadas neste trabalho.

$$AC = \mathbf{54,65612} + \mathbf{1,92950}\, M + 7{,}22438\, T_{es} + \mathbf{1,86603}\, T_{eo} + 0{,}04950\, S\,, com\ R^2 \cong 0{,}38 \quad (80)$$

Na Eq. 80, *AC* é ângulo de contato, *M* é o mineral (ponto de zero carga) e T_{es} é o tempo de envelhecimento em salmoura, T_{eo} é o tempo de envelhecimento em óleo e S é o solvente (constante dielétrica). Dos coeficientes, somente os das variáveis *M* e T_{eo} atingiram significância estatística com α = 5% e por isso estão em negrito. Observe que os coeficientes dessas duas variáveis são positivos. Portanto, quanto maior o ponto de zero carga do mineral (*i.e.* quanto mais próximo de valores característicos de carbonatos) e quanto maior o tempo de envelhecimento em óleo, maior o AC, conforme esperado. O resultado completo da regressão, incluindo os coeficientes e os intervalos de confiança de 95% para eles, encontra-se no Anexo I, na Figura 47.

Pelo R^2 obtido, conclui-se que aproximadamente 38% da variabilidade dos valores de ângulo de contato é explicada pelas variáveis independentes estudadas. Logo, o valor confirma que Fenômenos de Superfície apresentam não-linearidade. A Figura 38 apresenta a reta ajustada e os pontos experimentais.

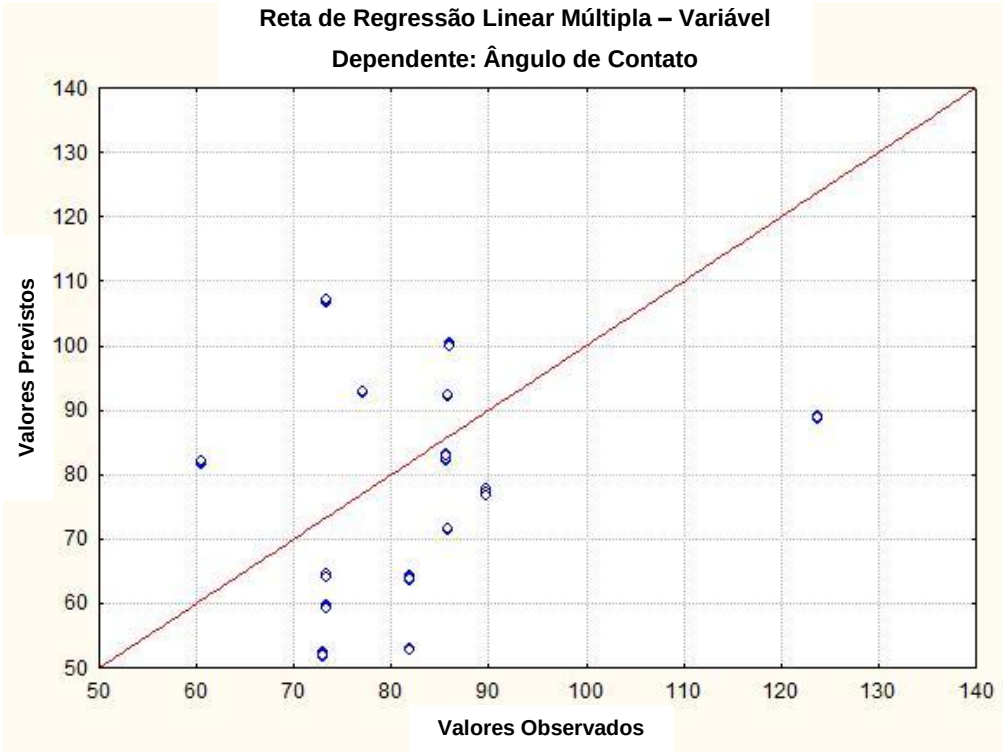

Figura 38 - Reta ajustada pela Regressão Linear Múltipla e medições experimentais (pontos).

Uma das diversas utilidades do R^2, caso ele seja baixo, é sinalizar que talvez seja necessário o aumento da complexidade do modelo[21] (*e.g.* acrescentando mais variáveis independentes, como a temperatura), para que este se torne um pouco mais representativo da realidade. De qualquer maneira, é muito importante ressaltar que o fato de a população de ACs apresentar não-linearidade não inviabiliza as análises realizadas: o resultado de qualquer Regressão Múltipla é o melhor modelo linear para um fenômeno real estudado e, dentro deste modelo, as variáveis independentes mais influentes são as encontradas pela ANOVA.

Retirando do modelo as variáveis independentes que não atingiram significância estatística na Regressão, *i.e.* solvente (constante dielétrica) e T_{es}, o novo R^2 obtido foi baixo e muito similar ao anterior: 0,37. Isso mais uma vez confirma a não-linearidade dos Fenômenos de Superfície e também que as variáveis excluídas são pouco influentes no resultado de ângulo de contato.

A última técnica empregada foi o Diagrama de Efeitos de Pareto. Para sua obtenção, os dados utilizados foram os mesmos da ANOVA, ou seja, as 320 medições

[21] Um modelo pode ser entendido como uma representação muito simplificada da realidade, intrinsecamente complexa. Como afirmou o estatístico George E. P. Box, “essencialmente, todos os modelos estão errados, mas alguns são úteis” [44].

que respeitavam a condição de homocedasticidade. A Figura 39 apresenta os resultados. Nela é possível observar novamente que a variável mais influente para os resultados de AC é o tempo de envelhecimento em óleo, com o mineral em segundo lugar (ambas com significância estatística, α = 5%). Assim como as análises anteriores haviam concluído, o solvente não influi no AC.

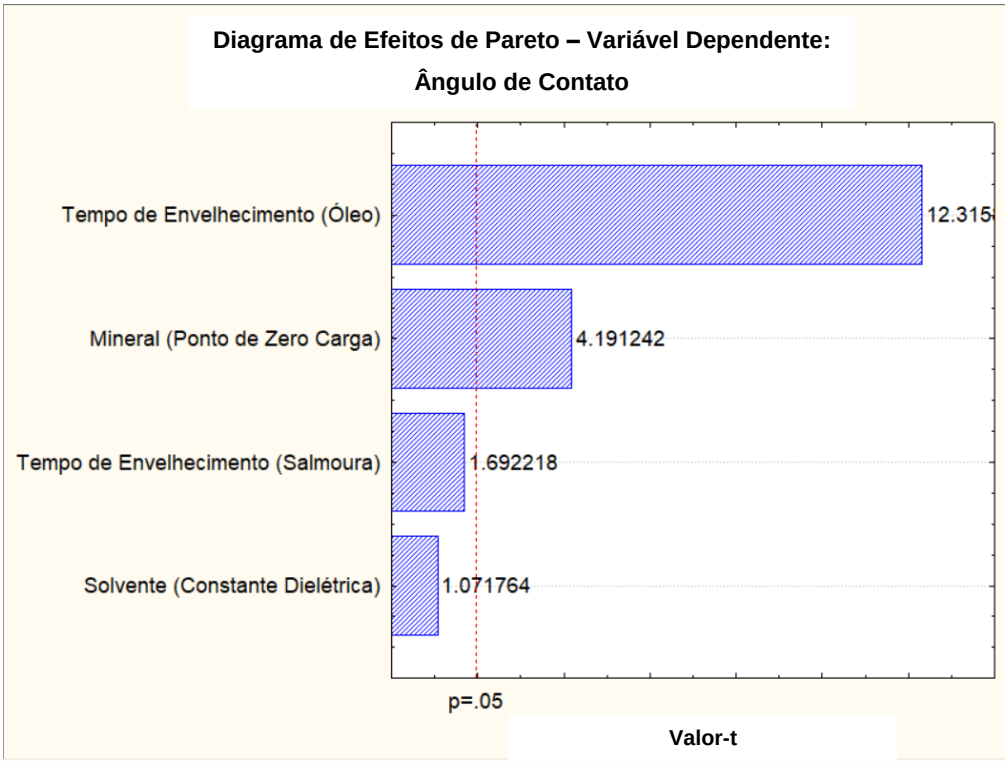

Figura 39 - Diagrama de Efeitos de Pareto, construído para α = 5%.

De todas as análises realizadas, nenhuma encontrou influência significativa do solvente. Intuitivamente, era esperada a existência de uma relação entre o AC e o solvente utilizado para limpeza da amostra de mineral. Mesmo com um SOP[22] para a limpeza, é a etapa dos procedimentos experimentais que envolve mais erros experimentais, visto que envolve maior ação humana. Portanto, é possível que o fator humano tenha interferido na limpeza das amostras a ponto de gerar diversas medições na faixa de molhabilidade neutra: 109 medições (34,06% das 320 medições) encontram-se neste intervalo. A média da população do total de 460 ângulos de contato também indica molhabilidade neutra (vide Tabela 18). Das 109 medições, 20 (18,35%) são com

[22] *Standard Operating Procedure*, ou na tradução livre para o Português, Procedimento Operacional Padrão. Um SOP pretende padronizar procedimentos experimentais para garantir a repetibilidade de medições e redução de erros.

nenhum solvente, 40 (36,70%) com salmoura, 20 (18,35%) com ciclohexano, 20 (18,35%) com hexadecano e 9 (8,26%) com tolueno. Somente o n-hexano não possui nenhuma medição nessa faixa. Logo, como há muitas medidas similares e pertencentes a um mesmo intervalo, as análises, que são fundamentadas em variações, não encontraram relevância no solvente a 5% de significância.

Quanto ao tempo de envelhecimento em salmoura, provou-se que não é influente em todas as análises exceto na matriz de correlações, que apresentou uma correlação fraca entre T_{es} e AC. Trata-se de um falso positivo: T_{es} varia apenas entre 0 e 1 dia. As medições com 0 dias, *i.e.* sem envelhecimento em salmoura e em óleo, constituem apenas 3 experimentos do total de 23 (já descontado o incidente no experimento 3). Em todos os outros, T_{es} = 1 dia e há envelhecimento em óleo, o que, como já foi descrito, aumenta o AC. Então, os 20 experimentos com envelhecimento em óleo e com maiores ACs comparados aos 3 experimentos sem envelhecimento e com menores ACs geraram a correlação para T_{es}. Portanto, na realidade a variação do ângulo de contato está relacionada ao T_{eo} e não ao T_{es}.

Finalmente, é muito importante destacar que 75% das análises estatísticas realizadas identificaram o T_{eo} e o mineral como principais fatores dos experimentos, atingindo significância estatística. Portanto, pode-se afirmar com 95% de confiança que os vínculos entre essas variáveis e os ACs não são meros frutos do acaso (aleatoriedade), mas sim que representam relações substanciais. Os resultados atingidos pelas análises são, portanto, muito robustos.

6. Considerações Finais

A literatura não possui atualmente muitos estudos e publicações sobre mediçôes de ângulo de contato aplicadas a cenários do pré-sal brasileiro, principalmente se consideradas a quantidade de artigos científicos que envolvem análises estatísticas além do trabalho experimental. Portanto, o presente trabalho pretende contribuir neste sentido.

Dado o escopo do apresentado na seção 1.2, resultados satisfatórios foram obtidos: centenas de medições experimentais de ângulo de contato foram realizadas e a análise estatística dos dados encontrou relações interessantes entre as variáveis independentes e os resultados de AC.

Os ângulos de contato obtidos neste trabalho podem ter diversas aplicações, como a obtenção da tensão de adesão dos minerais utilizados e a caracterização de suas molhabilidades em condições análogas às condições dos reservatórios do pré-sal brasileiro. Além disso, uma de suas principais aplicações é servir como valiosa variável de entrada (*input*) em simuladores computacionais de escoamento multifásico em meio poroso, que buscam obter previsões de produção e de fatores de recuperação de reservatórios.

Com relação ao estudo da molhabilidade dos minerais representativos das rochas reservatórios do pré-sal brasileiro, verificou-se que os minerais calcita e dolomita têm tendência a serem molháveis ao óleo e o quartzo a ser molhável à água, conforme DONALDSON E ALAM [1] afirmam.

Quanto aos resultados da análise estatística, verificou-se que as variáveis independentes que mais influenciam nos resultados de ângulo de contato são: em primeiro lugar, o tempo de envelhecimento em óleo, e, em segundo lugar, o mineral. Das quatro análises realizadas, três chegaram a esta conclusão. O fato de essas variáveis atingirem significância estatística é primordial, haja vista que é a confirmação matemática de que as relações obtidas não foram ao acaso (aleatórias): um maior tempo de envelhecimento em óleo e um maior ponto de zero carga (pzc) do mineral implicam em maiores ângulos de contato com 95% de confiança.

Diante do que foi apresentado na seção 5 e ao longo do trabalho, é interessante observar a diferença entre uma análise puramente qualitativa e uma quantitativa: é da natureza humana tentar encontrar padrões em números e na Natureza, porém somente poderosas ferramentas como a Estatística conseguem fundamentar interpretações e garantir suas neutralidades. Além disso, se a análise quantitativa não for de grande

escala e extremamente complexa, o custo de sua realização é muito baixo, especialmente se comparado ao valor de seus resultados.

Como trabalhos futuros, seria de grande valia a comparação dos dados experimentais obtidos neste projeto com dados de ACs utilizando a centrífuga como método de limpeza dos *rock slabs*. Com este equipamento, os erros humanos são minimizados e as incertezas na etapa de preparação de amostras seriam em maior parte consequências do desempenho da centrífuga, que pode ser considerado constante, dadas condições de operação satisfatórias. Além disso, enriqueceriam o trabalho: o aumento do número de medições para cada condição experimental, o que aumentaria a robustez e precisão das análises estatísticas; a inclusão de amostras de rochas reais do pré-sal; a utilização de fluidos vivos (*e.g.* com gás carbônico dissolvido) nos testes; e a mudança das condições termodinâmicas de temperatura e pressão. Os objetivos seriam aproximar cada vez mais os parâmetros experimentais e os fluidos das condições de reservatório do pré-sal brasileiro e tornar as análises as mais acuradas possíveis, tornando o trabalho ainda mais relevante.

7. Referências Bibliográficas

[1] DONALDSON, E. C. e ALAM, W. ***Wettability***, 1ª edição, 2008. Houston, Estados Unidos da América, Gulf Publishing Company.

[2] BRITISH PETROLEUM - BP [S.I.][s.n.]. **BP Energy Outlook**, 2017. Disponível em: <https://www.bp.com/content/dam/bp/pdf/energy-economics/energy-outlook-2017/bp-energy-outlook-2017.pdf>. Acesso em: 07 de janeiro de 2018, 18:43:00.

[3] YERGIN, D. ***The Prize: The Epic Quest for Oil, Money & Power***, 3ª edição, 2009. Nova Iorque, Estados Unidos da América, Simon & Schuster Inc.

[4] COSTA NETO, P. L. O. ***Estatística***, 2ª edição, 2002. São Paulo, Brasil, Edgard Blucher.

[5] SALSBURG, D. ***The Lady Tasting Tea (How Statistics Revolutionized Science in the Twentieth Century)***, 2ª edição, 2002. Nova Iorque, Estados Unidos da América, Henry Holt & Co.

[6] ALESSANDRO, J. [S.I.][s.n.]. "Aula 20 – Medidas de Assimetria", 2012. Disponível em: <https://pt.slideshare.net/JoaoAlessandro/aula-20-medidas-de-assimetria>. Acesso em: 08 de janeiro de 2018, 19:00:00.

[7] BARBAS, R. [S.I.][s.n.]. "Estatística", 2015. Disponível em: <http://slideplayer.com.br/slide/2324507/8/images/5/Curtose+e+o+coeficiente.jpg>. Acesso em: 08 de janeiro de 2018, 19:25:00.

[8] CASELLA, G. e BERGER, R. L. ***Statistical Inference***, 1ª edição, 1990. Belmont, Estados Unidos da América, Duxbury Press.

[9] SALDANHA, J. M. B. ***Estatística Aplicada I***, 2014. Notas de Aula, Rio de Janeiro, Brasil, Universidade Federal do Rio de Janeiro (UFRJ).

[10] NEYMAN, J. **Frequentist Probability and Frequentist Statistics**, 1977. "Synthese", 36(1), p. 97-131. Berkeley, Estados Unidos da América, D. Reidel Publishing Company.

[11] MEYER, P. L. ***Introductory Probability and Statistical Applications***, 2ª edição, 1970. Estados Unidos da América, Addison-Wesley Publishing Company Inc.

[12] HASLWANTER, T.. ***An Introduction to Statistics with Python***, 1ª edição, 2016. Zurique, Suíça, Springer International Publishing Switzerland.

[13] DEA GROUP [S.I.][s.n.]. ***Core Analysis***, 2018. Disponível em: <http://www.dea-group.com/de/node/862>. Acesso em: 11 de janeiro de 2018, 09:26:00.

[14] THE PENNSYLVANIA STATE UNIVERSITY [S.l.][s.n.]. ***STAT 414/415 – Probability Theory and Mathematical Statistics***, 2018. Disponível em: <https://onlinecourses.science.psu.edu/stat414/node/174>. Acesso em: 13 de janeiro de 2018, 08:55:00.

[15] PREL, J. D., HOMMEL, G., ROHRIG, B. E BLETTNER, M. ***Confidence Interval or P-Value? Part 4 of a Series on Evaluation of Scientific Publications***, 2009. "Deutsches Ärzteblatt international", 106(19), p. 335-339. Berlin, Alemanha, Deutsches Ärzteverlag GmbH.

[16] FISHER, R. A. ***Statistical Methods for Research Workers***, 1925. Edinburgo, Reino Unido, Oliver and Boyd.

[17] SAWILOWSKY, S. S., BLAIR, R. C., HIGGINS, J. J. ***An Investigation of the Type I Error and Power Properties of the Rank Transform Procedure in Factorial ANOVA***, 1989. "Journal of Educational Statistics", 14(3), p. 255-267. Estados Unidos da América, American Educational Research Association and American Statistical Association.

[18] SNEDECOR, G. W. e COCHRAN, W. G. ***Statistical Methods***, 8ª edição, 1989. Iowa, Estados Unidos da América, Iowa State University Press.

[19] PEARSON, K. ***On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it can be Reasonably Supposed to have Arisen from Random Sampling***, 1900. Londres, Inglaterra, University College of London.

[20] CHAKRAVARTI, I. M. ***Handbook of Methods of Applied Statistics***, Volume I, 1ª edição, 1967, p. 392-394. Nova Jersey, Estados Unidos da América, John Wiley & Sons.

[21] LACEY, M. [s.n.]. ***Statistical Topics***, 1998. Department of Statistics, Yale University. Disponível em: <http://www.stat.yale.edu/Courses/1997-98/101/stat101.htm>. Acesso em: 17 de janeiro de 2018, 15:30:00.

[22] BENT, M. ***The Pareto Principle on YouTube***, 2010. Disponível em: <http://referaat.cs.utwente.nl/conference/12/paper/6998/the-pareto-principle-on-youtube.pdf>. Enschede, Holanda, University of Twente.

[23] WILKINSON, L. ***Revising the Pareto Chart***, 2006. "The American Statistician", 60(4), p. 332-334. Londres, Inglaterra, Taylor & Francis Group.

[24] ANTONY, J. ***Design of Experiments for Engineers and Scientists***, 1ª edição, 2003. Estados Unidos da América, Elsevier Science.

[25] “A forma das gotas d’água”, 2011 [S.l.].[s.n.]. Disponível em: <http://umaquimicairresistivel.blogspot.com.br/2011/12/forma-das-gotas-de-agua.html>. Acesso em: 29 de março de 2017, 11:40:00.

[26] ATKINS, P. e DE PAULA, J. ***Physical Chemistry***, 9ª edição, 2010. Nova Iorque, Estados Unidos, W.H Freeman and Company.

[27] DREXLER, S. G. ***Métodos de Recuperação Avançada de Petróleo***, 2016. Rio de Janeiro, Brasil, Notas de Aula, Universidade Federal do Rio de Janeiro.

[28] ROSA, A., CARVALHO, R. e XAVIER, D. ***Engenharia de Reservatórios de Petróleo***, 1ª edição, 2006. Rio de Janeiro, Brasil, Editora Interciência.

[29] ROSEN, M. J. ***Surfactants and Interfacial Phenomena***, 3a edição, 2004. Hoboken, Estados Unidos da América, John Wiley & Sons Inc.

[30] KRAMER, M. R., EDELHOFF, D. e STAWARCZYK, B. ***Flexural Strength of Preheated Resin Composites and Bonding Properties to Glass-Ceramic and Dentin***, 2016. “Materials”, 9(2), p. 83. Basileia, Suíça, MDPI.

[31] CLACKAMAS COMMUNITY COLLEGE [S.l.][s.n.]. ***Introductory Chemistry***, 2011. Disponível em: <http://dl.clackamas.edu/ch104/lesson9polarity_simple.html>. Acesso em: 14 de janeiro de 2018, 13:04:00.

[32] BROWN, T. E. *et al.* ***Chemistry, The Central Science***, 9ª edição, 2002. Estados Unidos da América, Pearson.

[33] BRACCO, G. e HOLST, B. ***Surface Science Techniques***, 1ª edição, 2013. Berlin, Alemanha, Springer-Verlag Berlin Heidelberg.

[34] KRÜSS GmbH [s.n.]. ***DSA100 Manual***, 2014. Hamburgo, Alemanha.

[35] EXTRAND, C. W. ***Uncertainty in Contact Angle Measurements from the Tangent Method***, 2016. “Journal of Adhesion Science and Technology”, 30(15). Londres, Inglaterra, Taylor & Francis Group.

[36] FAÇANHA, J. M. F., SOHRABI, M., DREXLER, S. G., COUTO, P. ***Experimental Investigation of Wettability for Brazilian Pre-Salt Carbonates***, 2016. “Rio Oil & Gas Conference”, IBP1455_16. Rio de Janeiro, Brazil.

[37] KOSMULSKI, M. ***The pH-Dependent Surface Charging and the Points of Zero Charge***, 2002. “Journal of Colloid and Interface Science”, nº 253, p. 77-87. Estados Unidos da América, Elsevier Science.

[38] AL SHALABI, E. W. e SEPEHRNOORI, K. ***Low Salinity and Engineered Water Injection for Sandstone and Carbonate Reservoirs***, 1ª edição, 2017. Oxford, Reino Unido, Elsevier Inc.

[39] HECTOR, L. G. e SCHULTZ, H. L. ***The Dielectric Constant of Air at Radiofrequencies***, 1936. "Physics", 7(4), p. 133-136. American Institute of Physics, Estados Unidos da América.

[40] LIDE, D. R. ***CRC Handbook of Chemistry and Physics***, 86ª edição, 2005. Boca Raton, Estados Unidos da América, CRC Press.

[41] MARYOTT, A. A. e SMITH, E. R. ***Table of Dielectric Constants of Pure Liquids***, 1951. "National Bureau of Standards Circular", nº 514, Washington, D. C., Estados Unidos da América.

[42] ARALAGUPPI, M. I. *et al.* ***Thermodynamic Interactions in Mixtures of Bromoform with Hydrocarbons***, 1991. "Journal of Physical Chemistry", nº 95, p. 5299-5308. American Chemical Society, Estados Unidos da América.

[43] ZAIONTZ, C. ***Assumptions for ANOVA***, 2017. Disponível em: <http://www.real-statistics.com/one-way-analysis-of-variance-anova/assumptions-anova/>. Acessado em: 17 de janeiro de 2018, 20:54:00

[44] BOX, G. E. P. ***Science and Statistics***, 1976. "Journal of the American Statistical Association", nº 71, p. 791-799. American Statistical Society, Estados Unidos da América.

Anexo I – Statsoft Statistica 7.0

i. Exemplo da seção 2.3.4 (Empresa de Geofísica):

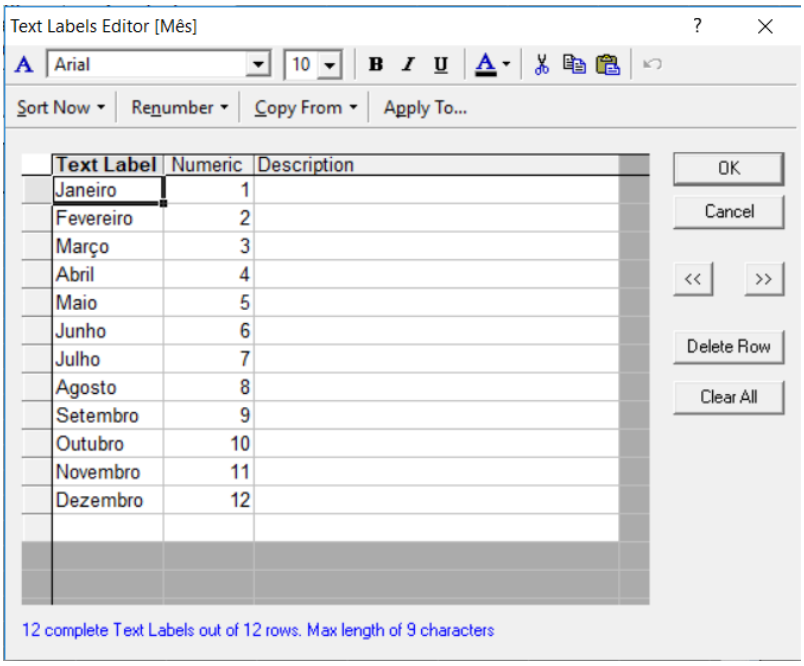

Figura 40 - Geração de rótulos e seus equivalentes numéricos para o *input* no *software*.

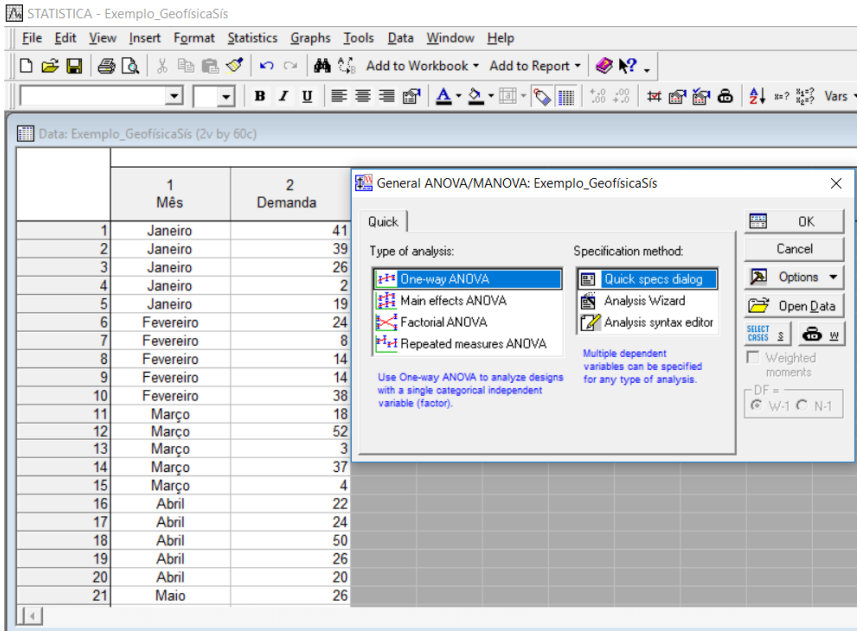

Figura 41 - *Inputs* para a *One-Way ANOVA* realizada na seção 2.3.4.

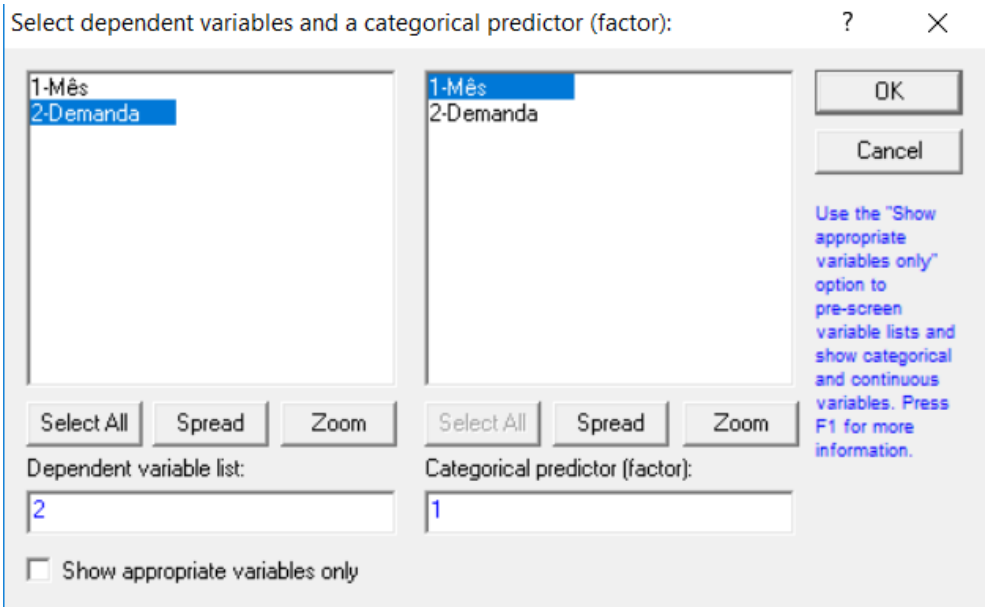

Figura 42 - Definição da variável dependente e independente.

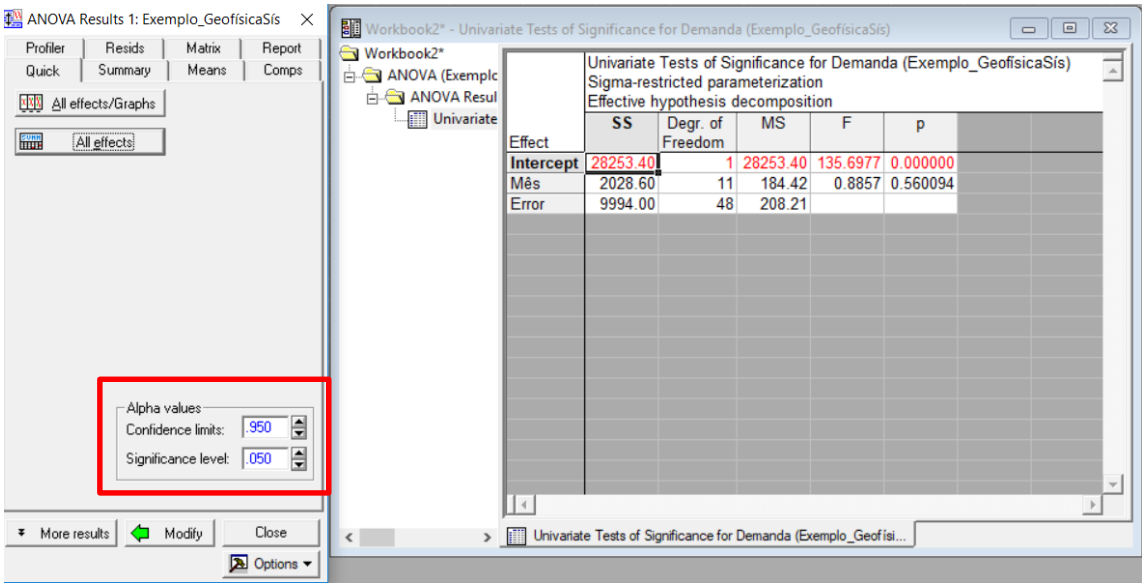

Figura 43 - Resultado da ANOVA. Repare no nível de significância utilizado (em destaque): 5%.

ii. Análise estatítica da seção 5:

Data: ANOVA_PG_FINAL (5v by 480c)

	1 Mineral (Ponto de Zero Carga)	2 Tempo de Envelhecimento (Salmoura)	3 Tempo de Envelhecimento (Óleo)	4 Solvente (Constante Dielétrica)	5 **Ângulo de Contato**
1	3	0	0	1	81.5
2	3	0	0	1	81.6
3	3	0	0	1	81.6
4	3	0	0	1	81.8
5	3	0	0	1	81.6
6	3	0	0	1	81.5
7	3	0	0	1	81.6
8	3	0	0	1	81.5
9	3	0	0	1	81.7
10	3	0	0	1	81.5
11	3	0	0	1	81.7
12	3	0	0	1	81.5
13	3	0	0	1	81.6
14	3	0	0	1	81.7
15	3	0	0	1	81.5
16	3	0	0	1	81.7
17	3	0	0	1	81.7
18	3	0	0	1	82.1
19	3	0	0	1	81.8
20	3	0	0	1	82.1
21	3	1	3	78.54	92.8
22	3	1	3	78.54	92.7
23	3	1	3	78.54	92.7
24	3	1	3	78.54	92.7
25	3	1	3	78.54	92.7
26	3	1	3	78.54	92.9
27	3	1	3	78.54	92.8
28	3	1	3	78.54	92.9
29	3	1	3	78.54	92.7

Figura 44 - Parte da tabela de *input*.

Data: ANOVA_PG_FINAL (5v by 480c)

	1 Mineral (Ponto de Zero Carga)	2 Tempo de Envelhecimento (Salmoura)	3 Tempo de Envelhecimento (Óleo)	4 **Solvente (Constante Dielétrica)**	5 Ângulo de Contato
15	3	1	3	2.379	59.4
16	3	1	3	2.379	59.4
17	3	1	3	2.379	59.4
18	3	1	3	2.379	59.4
19	3	1	3	2.379	59.4
20	3	1	3	2.379	59.3
21	3	1	15	2.015	143.8
22	3	1	15	2.015	144.6
23	3	1	15	2.015	143.6
24	3	1	15	2.015	144.9
25	3	1	15	2.015	145
26	3	1	15	2.015	143.7
27	3	1	15	2.015	143.7
28	3	1	15	2.015	143.7
29	3	1	15	2.015	142.9
30	3	1	15	2.015	143.1
31	3	1	15	2.015	143.7
32	3	1	15	2.015	143.9
33	3	1	15	2.015	143.2
34	3	1	15	2.015	143.9
35	3	1	15	2.015	143.7
36	3	1	15	2.015	143.2
37	3	1	15	2.015	143.1
38	3	1	15	2.015	143.4
39	3	1	15	2.015	144.1
40	3	1	15	2.015	143.3
41	3	1	30	2.015	89.1
42	3	1	30	2.015	88.6
43	3	1	30	2.015	89.1

Figura 45 – Exemplo de exclusão de experimentos (subpopulações) cujas variâncias não respeitavam a regra prática da ANOVA.

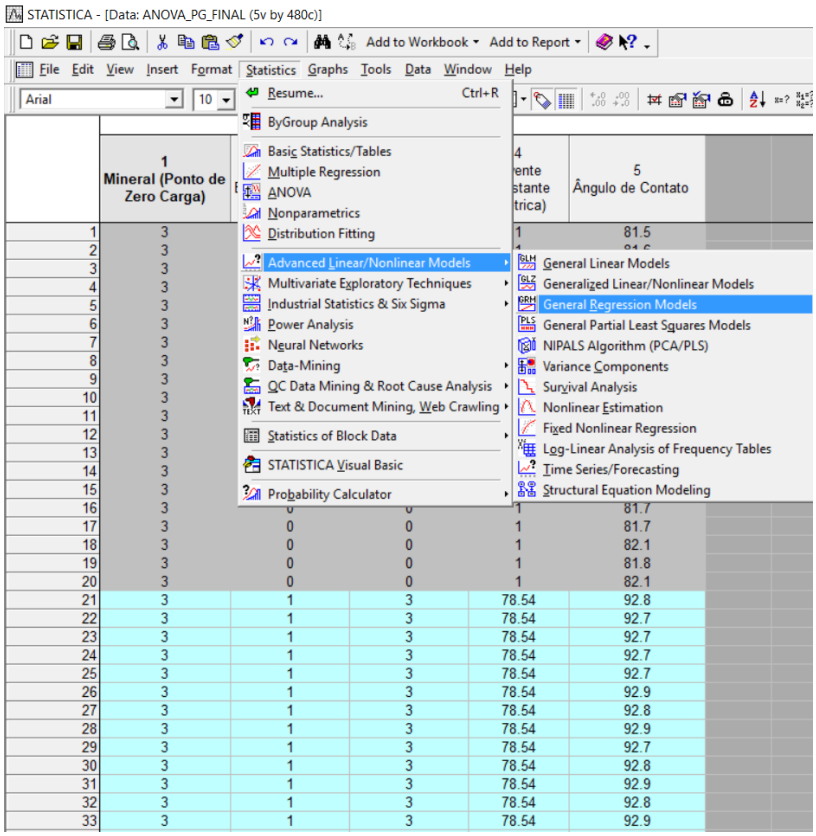

Figura 46 – Iniciando a Regressão Linear Múltipla e ANOVA por Regressão Linear Múltipla.

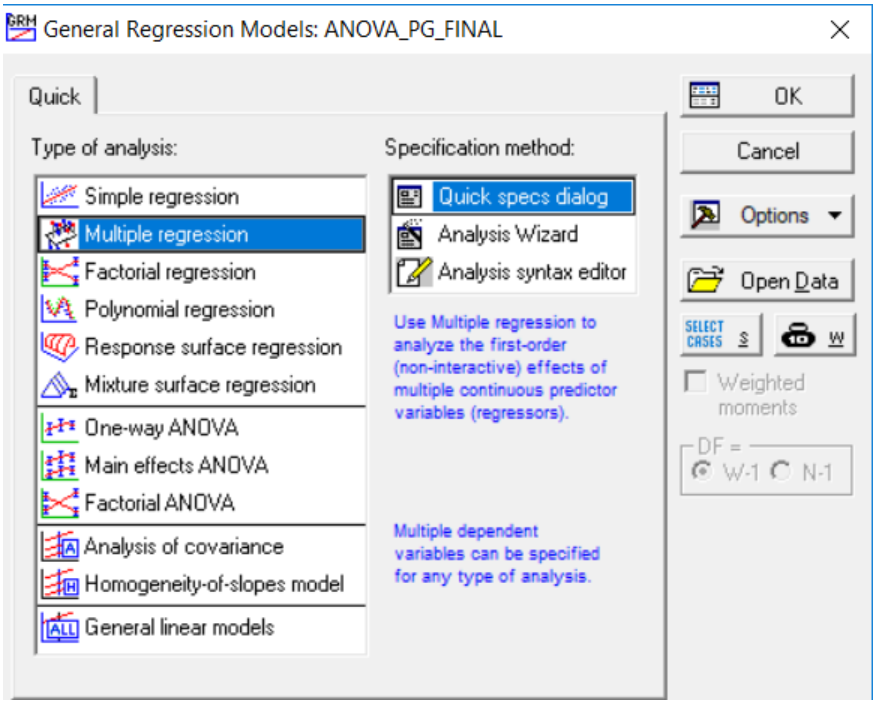

Figura 47 - Seleção da Regressão Linear Múltipla.

Effect	Parameter Estimates (ANOVA_PG_FINAL) Sigma-restricted parameterization Ângulo de Contato Param.	Ângulo de Contato Std.Err	Ângulo de Contato t	Ângulo de Contato p	-95.00% Cnf.Lmt	+95.00% Cnf.Lmt	Ângulo de Contato Beta (ß)	Ângulo de Contato St.Err.ß	-95.00% Cnf.Lmt	+95.00% Cnf.Lmt
Intercept	54.65612	4.688930	11.65642	0.000000	45.43054	63.88171				
Mineral (Ponto de Zero Carga)	1.92950	0.460365	4.19124	0.000036	1.02372	2.83528	0.185786	0.044327	0.098571	0.273001
Tempo de Envelhecimento (Salmoura)	7.22438	4.269177	1.69222	0.091593	-1.17533	15.62408	0.080486	0.047563	-0.013094	0.174066
Tempo de Envelhecimento (Óleo)	1.86603	0.151514	12.31588	0.000000	1.56792	2.16414	0.586958	0.047659	0.493189	0.680728
Solvente (Constante Dielétrica)	0.04950	0.046182	1.07176	0.284646	-0.04137	0.14036	0.049853	0.046515	-0.041666	0.141371

Figura 48 - Resultado completo da Regressão Linear Múltipla.

Anexo II – Distribuições Tabeladas

GL	V1																			
V2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20	40	60	120	240
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.0	243.9	244.7	245.4	245.9	248.0	251.1	252.2	253.3	253.8
2	18.513	19.000	19.164	19.247	19.296	19.329	19.353	19.371	19.385	19.396	19.405	19.412	19.419	19.424	19.429	19.446	19.471	19.479	19.487	19.492
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.785	8.763	8.745	8.729	8.715	8.703	8.660	8.594	8.572	8.549	8.538
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964	5.936	5.912	5.891	5.873	5.858	5.803	5.717	5.688	5.658	5.643
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.704	4.678	4.655	4.636	4.619	4.558	4.464	4.431	4.398	4.382
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060	4.027	4.000	3.976	3.956	3.938	3.874	3.774	3.740	3.705	3.687
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637	3.603	3.575	3.550	3.529	3.511	3.445	3.340	3.304	3.267	3.249
8	5.318	4.459	4.066	3.838	3.688	3.581	3.500	3.438	3.388	3.347	3.313	3.284	3.259	3.237	3.218	3.150	3.043	3.005	2.967	2.947
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137	3.102	3.073	3.048	3.025	3.006	2.936	2.826	2.787	2.748	2.727
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.943	2.913	2.887	2.865	2.845	2.774	2.661	2.621	2.580	2.559
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854	2.818	2.788	2.761	2.739	2.719	2.646	2.531	2.490	2.448	2.426
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753	2.717	2.687	2.660	2.637	2.617	2.544	2.426	2.384	2.341	2.319
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671	2.635	2.604	2.577	2.554	2.533	2.459	2.339	2.297	2.252	2.230
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602	2.565	2.534	2.507	2.484	2.463	2.388	2.266	2.223	2.178	2.155
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.507	2.475	2.448	2.424	2.403	2.328	2.204	2.160	2.114	2.090
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494	2.456	2.425	2.397	2.373	2.352	2.276	2.151	2.106	2.059	2.035
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450	2.413	2.381	2.353	2.329	2.308	2.230	2.104	2.058	2.011	1.986
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412	2.374	2.342	2.314	2.290	2.269	2.191	2.063	2.017	1.968	1.943
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378	2.340	2.308	2.280	2.256	2.234	2.155	2.026	1.980	1.930	1.905
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348	2.310	2.278	2.250	2.225	2.203	2.124	1.994	1.946	1.896	1.870
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321	2.283	2.250	2.222	2.197	2.176	2.096	1.965	1.916	1.866	1.839
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297	2.259	2.226	2.198	2.173	2.151	2.071	1.938	1.889	1.838	1.811
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275	2.236	2.204	2.175	2.150	2.128	2.048	1.914	1.865	1.813	1.785
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255	2.216	2.183	2.155	2.130	2.108	2.027	1.892	1.842	1.790	1.762
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236	2.198	2.165	2.136	2.111	2.089	2.007	1.872	1.822	1.768	1.740
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220	2.181	2.148	2.119	2.094	2.072	1.990	1.853	1.803	1.749	1.720
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204	2.166	2.132	2.103	2.078	2.056	1.974	1.836	1.785	1.731	1.702
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190	2.151	2.118	2.089	2.064	2.041	1.959	1.820	1.769	1.714	1.685
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177	2.138	2.104	2.075	2.050	2.027	1.945	1.806	1.754	1.698	1.669
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165	2.126	2.092	2.063	2.037	2.015	1.932	1.792	1.740	1.683	1.654
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077	2.038	2.003	1.974	1.948	1.924	1.839	1.693	1.637	1.577	1.544
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130	2.073	2.026	1.986	1.952	1.921	1.895	1.871	1.784	1.634	1.576	1.511	1.476
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993	1.952	1.917	1.887	1.860	1.836	1.748	1.594	1.534	1.467	1.430
80	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056	1.999	1.951	1.910	1.875	1.845	1.817	1.793	1.703	1.545	1.482	1.411	1.370
100	3.936	3.087	2.696	2.463	2.305	2.191	2.103	2.032	1.975	1.927	1.886	1.850	1.819	1.792	1.768	1.676	1.515	1.450	1.376	1.333
120	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016	1.959	1.910	1.869	1.834	1.803	1.775	1.750	1.659	1.495	1.429	1.352	1.307
240	3.881	3.033	2.642	2.409	2.252	2.136	2.048	1.977	1.919	1.870	1.829	1.793	1.761	1.733	1.708	1.614	1.445	1.375	1.290	1.237

Figura 49 - Valores tabelados da distribuição F de Snedecor considerando um nível de significância de 5%. V1 são os graus de liberdade no numerador e V2 os graus de liberdade no denominador. Fonte: [9].

Anexo III – Medições Experimentais

Tabela 23 - Dados experimentais obtidos (20 medidas estáveis para cada experimento).

Mineral	Tempo de Envelhecimento em Salmoura [dias]	Tempo de Envelhecimento em Óleo [dias]	Solvente	Ângulo de Contato (º)
Quartzo	0	0	Nenhum	81,5
Quartzo	0	0	Nenhum	81,6
Quartzo	0	0	Nenhum	81,6
Quartzo	0	0	Nenhum	81,8
Quartzo	0	0	Nenhum	81,6
Quartzo	0	0	Nenhum	81,5
Quartzo	0	0	Nenhum	81,6
Quartzo	0	0	Nenhum	81,5
Quartzo	0	0	Nenhum	81,7
Quartzo	0	0	Nenhum	81,5
Quartzo	0	0	Nenhum	81,7
Quartzo	0	0	Nenhum	81,5
Quartzo	0	0	Nenhum	81,6
Quartzo	0	0	Nenhum	81,7
Quartzo	0	0	Nenhum	81,5
Quartzo	0	0	Nenhum	81,7
Quartzo	0	0	Nenhum	81,7
Quartzo	0	0	Nenhum	82,1
Quartzo	0	0	Nenhum	81,8
Quartzo	0	0	Nenhum	82,1
Quartzo	1	3	Salmoura	92,8
Quartzo	1	3	Salmoura	92,7
Quartzo	1	3	Salmoura	92,7
Quartzo	1	3	Salmoura	92,7
Quartzo	1	3	Salmoura	92,7
Quartzo	1	3	Salmoura	92,9
Quartzo	1	3	Salmoura	92,8
Quartzo	1	3	Salmoura	92,9
Quartzo	1	3	Salmoura	92,7
Quartzo	1	3	Salmoura	92,8
Quartzo	1	3	Salmoura	92,9
Quartzo	1	3	Salmoura	92,8
Quartzo	1	3	Salmoura	92,9
Quartzo	1	3	Salmoura	92,8
Quartzo	1	3	Salmoura	92,7
Quartzo	1	3	Salmoura	92,8
Quartzo	1	3	Salmoura	92,9
Quartzo	1	3	Salmoura	92,9

Quartzo	1	3	Salmoura	92,8
Quartzo	1	3	Salmoura	92,8
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	N-Hexano	
Quartzo	1	3	Ciclohexano	64,4
Quartzo	1	3	Ciclohexano	64,5
Quartzo	1	3	Ciclohexano	64,6
Quartzo	1	3	Ciclohexano	64,5
Quartzo	1	3	Ciclohexano	64,4
Quartzo	1	3	Ciclohexano	64,5
Quartzo	1	3	Ciclohexano	64,4
Quartzo	1	3	Ciclohexano	64,4
Quartzo	1	3	Ciclohexano	64,6
Quartzo	1	3	Ciclohexano	64,3
Quartzo	1	3	Ciclohexano	64,5
Quartzo	1	3	Ciclohexano	64,4
Quartzo	1	3	Ciclohexano	64,5
Quartzo	1	3	Ciclohexano	64,2
Quartzo	1	3	Ciclohexano	64,2
Quartzo	1	3	Ciclohexano	64,5
Quartzo	1	3	Ciclohexano	64,7
Quartzo	1	3	Ciclohexano	64,3
Quartzo	1	3	Ciclohexano	64,5
Quartzo	1	3	Ciclohexano	64,1
Quartzo	1	3	N-Hexadecano	106,9
Quartzo	1	3	N-Hexadecano	106,9
Quartzo	1	3	N-Hexadecano	106,8
Quartzo	1	3	N-Hexadecano	107,1

Quartzo	1	3	N-Hexadecano	107,2
Quartzo	1	3	N-Hexadecano	106,8
Quartzo	1	3	N-Hexadecano	107
Quartzo	1	3	N-Hexadecano	106,7
Quartzo	1	3	N-Hexadecano	107,1
Quartzo	1	3	N-Hexadecano	107,1
Quartzo	1	3	N-Hexadecano	107
Quartzo	1	3	N-Hexadecano	107,1
Quartzo	1	3	N-Hexadecano	107
Quartzo	1	3	N-Hexadecano	107,2
Quartzo	1	3	N-Hexadecano	107
Quartzo	1	3	N-Hexadecano	106,9
Quartzo	1	3	N-Hexadecano	107,2
Quartzo	1	3	N-Hexadecano	107,1
Quartzo	1	3	N-Hexadecano	107,1
Quartzo	1	3	N-Hexadecano	107,2
Quartzo	1	3	Tolueno	59,6
Quartzo	1	3	Tolueno	59,7
Quartzo	1	3	Tolueno	59,5
Quartzo	1	3	Tolueno	59,5
Quartzo	1	3	Tolueno	59,5
Quartzo	1	3	Tolueno	59,4
Quartzo	1	3	Tolueno	59,6
Quartzo	1	3	Tolueno	59,5
Quartzo	1	3	Tolueno	59,4
Quartzo	1	3	Tolueno	59,4
Quartzo	1	3	Tolueno	59,5
Quartzo	1	3	Tolueno	59,5
Quartzo	1	3	Tolueno	59,5
Quartzo	1	3	Tolueno	59,5
Quartzo	1	3	Tolueno	59,4
Quartzo	1	3	Tolueno	59,4
Quartzo	1	3	Tolueno	59,4
Quartzo	1	3	Tolueno	59,4
Quartzo	1	3	Tolueno	59,4
Quartzo	1	3	Tolueno	59,3
Quartzo	1	15	Ciclohexano	143,8
Quartzo	1	15	Ciclohexano	144,6
Quartzo	1	15	Ciclohexano	143,6
Quartzo	1	15	Ciclohexano	144,9
Quartzo	1	15	Ciclohexano	145
Quartzo	1	15	Ciclohexano	143,7
Quartzo	1	15	Ciclohexano	143,7
Quartzo	1	15	Ciclohexano	143,7
Quartzo	1	15	Ciclohexano	142,9
Quartzo	1	15	Ciclohexano	143,1

Quartzo	1	15	Ciclohexano	143,7
Quartzo	1	15	Ciclohexano	143,9
Quartzo	1	15	Ciclohexano	143,2
Quartzo	1	15	Ciclohexano	143,9
Quartzo	1	15	Ciclohexano	143,7
Quartzo	1	15	Ciclohexano	143,2
Quartzo	1	15	Ciclohexano	143,1
Quartzo	1	15	Ciclohexano	143,4
Quartzo	1	15	Ciclohexano	144,1
Quartzo	1	15	Ciclohexano	143,3
Quartzo	1	30	Ciclohexano	89,1
Quartzo	1	30	Ciclohexano	88,6
Quartzo	1	30	Ciclohexano	89,1
Quartzo	1	30	Ciclohexano	89,1
Quartzo	1	30	Ciclohexano	89,1
Quartzo	1	30	Ciclohexano	89,1
Quartzo	1	30	Ciclohexano	89,1
Quartzo	1	30	Ciclohexano	88,9
Quartzo	1	30	Ciclohexano	89
Quartzo	1	30	Ciclohexano	89,2
Quartzo	1	30	Ciclohexano	88,8
Quartzo	1	30	Ciclohexano	89
Quartzo	1	30	Ciclohexano	89
Quartzo	1	30	Ciclohexano	88,9
Quartzo	1	30	Ciclohexano	88,9
Quartzo	1	30	Ciclohexano	88,8
Quartzo	1	30	Ciclohexano	88,7
Quartzo	1	30	Ciclohexano	88,6
Quartzo	1	30	Ciclohexano	88,6
Quartzo	1	30	Ciclohexano	88,7
Calcita	0	0	Nenhum	51,7
Calcita	0	0	Nenhum	51,7
Calcita	0	0	Nenhum	51,8
Calcita	0	0	Nenhum	51,7
Calcita	0	0	Nenhum	51,7
Calcita	0	0	Nenhum	51,9
Calcita	0	0	Nenhum	51,7
Calcita	0	0	Nenhum	51,7
Calcita	0	0	Nenhum	52,3
Calcita	0	0	Nenhum	51,6
Calcita	0	0	Nenhum	51,7
Calcita	0	0	Nenhum	51,8
Calcita	0	0	Nenhum	51,8
Calcita	0	0	Nenhum	52,4
Calcita	0	0	Nenhum	51,8
Calcita	0	0	Nenhum	51,9

Calcita	0	0	Nenhum	52
Calcita	0	0	Nenhum	52,2
Calcita	0	0	Nenhum	52
Calcita	0	0	Nenhum	51,9
Calcita	1	3	Salmoura	77,2
Calcita	1	3	Salmoura	77
Calcita	1	3	Salmoura	77
Calcita	1	3	Salmoura	76,8
Calcita	1	3	Salmoura	77
Calcita	1	3	Salmoura	77,3
Calcita	1	3	Salmoura	76,9
Calcita	1	3	Salmoura	76,8
Calcita	1	3	Salmoura	77,4
Calcita	1	3	Salmoura	77,3
Calcita	1	3	Salmoura	77,8
Calcita	1	3	Salmoura	77,5
Calcita	1	3	Salmoura	77,7
Calcita	1	3	Salmoura	77,2
Calcita	1	3	Salmoura	77,1
Calcita	1	3	Salmoura	76,7
Calcita	1	3	Salmoura	77,1
Calcita	1	3	Salmoura	77,2
Calcita	1	3	Salmoura	77,3
Calcita	1	3	Salmoura	76,8
Calcita	1	3	N-Hexano	71,4
Calcita	1	3	N-Hexano	71,4
Calcita	1	3	N-Hexano	71,6
Calcita	1	3	N-Hexano	71,7
Calcita	1	3	N-Hexano	71,3
Calcita	1	3	N-Hexano	71,2
Calcita	1	3	N-Hexano	71,5
Calcita	1	3	N-Hexano	71,5
Calcita	1	3	N-Hexano	71,3
Calcita	1	3	N-Hexano	71,4
Calcita	1	3	N-Hexano	71,5
Calcita	1	3	N-Hexano	71,3
Calcita	1	3	N-Hexano	71,5
Calcita	1	3	N-Hexano	71,4
Calcita	1	3	N-Hexano	71,6
Calcita	1	3	N-Hexano	71,5
Calcita	1	3	N-Hexano	71,6
Calcita	1	3	N-Hexano	71,4
Calcita	1	3	N-Hexano	71,5
Calcita	1	3	N-Hexano	71,5
Calcita	1	3	Ciclohexano	62
Calcita	1	3	Ciclohexano	61,6

Calcita	1	3	Ciclohexano	60,8
Calcita	1	3	Ciclohexano	61,8
Calcita	1	3	Ciclohexano	62,3
Calcita	1	3	Ciclohexano	61,6
Calcita	1	3	Ciclohexano	59
Calcita	1	3	Ciclohexano	62,1
Calcita	1	3	Ciclohexano	62,1
Calcita	1	3	Ciclohexano	62
Calcita	1	3	Ciclohexano	62,3
Calcita	1	3	Ciclohexano	61,6
Calcita	1	3	Ciclohexano	61,8
Calcita	1	3	Ciclohexano	62
Calcita	1	3	Ciclohexano	61,3
Calcita	1	3	Ciclohexano	61,3
Calcita	1	3	Ciclohexano	60,5
Calcita	1	3	Ciclohexano	60,6
Calcita	1	3	Ciclohexano	59,9
Calcita	1	3	Ciclohexano	60,6
Calcita	1	3	N-Hexadecano	92
Calcita	1	3	N-Hexadecano	92
Calcita	1	3	N-Hexadecano	92,2
Calcita	1	3	N-Hexadecano	92
Calcita	1	3	N-Hexadecano	92,1
Calcita	1	3	N-Hexadecano	92,1
Calcita	1	3	N-Hexadecano	92
Calcita	1	3	N-Hexadecano	92,1
Calcita	1	3	N-Hexadecano	92,2
Calcita	1	3	N-Hexadecano	92,1
Calcita	1	3	N-Hexadecano	92,2
Calcita	1	3	N-Hexadecano	92,4
Calcita	1	3	N-Hexadecano	92,2
Calcita	1	3	N-Hexadecano	92,3
Calcita	1	3	N-Hexadecano	92,3
Calcita	1	3	N-Hexadecano	92,3
Calcita	1	3	N-Hexadecano	92,2
Calcita	1	3	N-Hexadecano	92,4
Calcita	1	3	N-Hexadecano	92,3
Calcita	1	3	N-Hexadecano	92,3
Calcita	1	3	Tolueno	100
Calcita	1	3	Tolueno	100,2
Calcita	1	3	Tolueno	99,9
Calcita	1	3	Tolueno	100
Calcita	1	3	Tolueno	100,1
Calcita	1	3	Tolueno	100
Calcita	1	3	Tolueno	100,1
Calcita	1	3	Tolueno	100,1

Calcita	1	3	Tolueno	100
Calcita	1	3	Tolueno	100,2
Calcita	1	3	Tolueno	100,2
Calcita	1	3	Tolueno	100
Calcita	1	3	Tolueno	100,1
Calcita	1	3	Tolueno	99,9
Calcita	1	3	Tolueno	100,2
Calcita	1	3	Tolueno	100,1
Calcita	1	3	Tolueno	100
Calcita	1	3	Tolueno	100,3
Calcita	1	3	Tolueno	100,1
Calcita	1	3	Tolueno	100
Calcita	1	15	Ciclohexano	162,2
Calcita	1	15	Ciclohexano	162,6
Calcita	1	15	Ciclohexano	163,2
Calcita	1	15	Ciclohexano	162,3
Calcita	1	15	Ciclohexano	162,9
Calcita	1	15	Ciclohexano	163,1
Calcita	1	15	Ciclohexano	162,8
Calcita	1	15	Ciclohexano	162,6
Calcita	1	15	Ciclohexano	162,9
Calcita	1	15	Ciclohexano	163,1
Calcita	1	15	Ciclohexano	163,2
Calcita	1	15	Ciclohexano	162,9
Calcita	1	15	Ciclohexano	163,2
Calcita	1	15	Ciclohexano	163
Calcita	1	15	Ciclohexano	163,1
Calcita	1	15	Ciclohexano	162,5
Calcita	1	15	Ciclohexano	162,9
Calcita	1	15	Ciclohexano	162,7
Calcita	1	15	Ciclohexano	162,6
Calcita	1	15	Ciclohexano	162,3
Calcita	1	30	Ciclohexano	112,6
Calcita	1	30	Ciclohexano	113,8
Calcita	1	30	Ciclohexano	114,9
Calcita	1	30	Ciclohexano	114,3
Calcita	1	30	Ciclohexano	114,3
Calcita	1	30	Ciclohexano	113,6
Calcita	1	30	Ciclohexano	113,2
Calcita	1	30	Ciclohexano	114
Calcita	1	30	Ciclohexano	113,4
Calcita	1	30	Ciclohexano	114,1
Calcita	1	30	Ciclohexano	115
Calcita	1	30	Ciclohexano	112,6
Calcita	1	30	Ciclohexano	113,9
Calcita	1	30	Ciclohexano	113,5

Calcita	1	30	Ciclohexano	113,1
Calcita	1	30	Ciclohexano	113,6
Calcita	1	30	Ciclohexano	111,6
Calcita	1	30	Ciclohexano	113,7
Calcita	1	30	Ciclohexano	113,2
Calcita	1	30	Ciclohexano	113,1
Dolomita	0	0	Nenhum	49,5
Dolomita	0	0	Nenhum	50,1
Dolomita	0	0	Nenhum	49,9
Dolomita	0	0	Nenhum	49,8
Dolomita	0	0	Nenhum	49,6
Dolomita	0	0	Nenhum	49
Dolomita	0	0	Nenhum	51,4
Dolomita	0	0	Nenhum	50,5
Dolomita	0	0	Nenhum	49,8
Dolomita	0	0	Nenhum	50,6
Dolomita	0	0	Nenhum	50,7
Dolomita	0	0	Nenhum	50,6
Dolomita	0	0	Nenhum	50,4
Dolomita	0	0	Nenhum	50,5
Dolomita	0	0	Nenhum	50,6
Dolomita	0	0	Nenhum	50,5
Dolomita	0	0	Nenhum	50,7
Dolomita	0	0	Nenhum	50,6
Dolomita	0	0	Nenhum	50,6
Dolomita	0	0	Nenhum	50,6
Dolomita	1	3	Salmoura	82,5
Dolomita	1	3	Salmoura	82,3
Dolomita	1	3	Salmoura	82,2
Dolomita	1	3	Salmoura	82,6
Dolomita	1	3	Salmoura	82,5
Dolomita	1	3	Salmoura	82,3
Dolomita	1	3	Salmoura	82,4
Dolomita	1	3	Salmoura	82,5
Dolomita	1	3	Salmoura	82,8
Dolomita	1	3	Salmoura	82,4
Dolomita	1	3	Salmoura	82,9
Dolomita	1	3	Salmoura	82,6
Dolomita	1	3	Salmoura	82,4
Dolomita	1	3	Salmoura	82,7
Dolomita	1	3	Salmoura	83,1
Dolomita	1	3	Salmoura	82,4
Dolomita	1	3	Salmoura	82,9
Dolomita	1	3	Salmoura	82,8
Dolomita	1	3	Salmoura	83,1
Dolomita	1	3	Salmoura	82,9

Dolomita	1	3	N-Hexano	128
Dolomita	1	3	N-Hexano	128,8
Dolomita	1	3	N-Hexano	128,6
Dolomita	1	3	N-Hexano	128,3
Dolomita	1	3	N-Hexano	128,2
Dolomita	1	3	N-Hexano	128,8
Dolomita	1	3	N-Hexano	129,2
Dolomita	1	3	N-Hexano	129
Dolomita	1	3	N-Hexano	129,1
Dolomita	1	3	N-Hexano	128,6
Dolomita	1	3	N-Hexano	129,2
Dolomita	1	3	N-Hexano	128,5
Dolomita	1	3	N-Hexano	129,3
Dolomita	1	3	N-Hexano	128,9
Dolomita	1	3	N-Hexano	129,5
Dolomita	1	3	N-Hexano	129,1
Dolomita	1	3	N-Hexano	129,8
Dolomita	1	3	N-Hexano	129,4
Dolomita	1	3	N-Hexano	129,8
Dolomita	1	3	N-Hexano	129,9
Dolomita	1	3	Ciclohexano	63,8
Dolomita	1	3	Ciclohexano	63,5
Dolomita	1	3	Ciclohexano	64,3
Dolomita	1	3	Ciclohexano	64,3
Dolomita	1	3	Ciclohexano	64,2
Dolomita	1	3	Ciclohexano	63,9
Dolomita	1	3	Ciclohexano	63,9
Dolomita	1	3	Ciclohexano	64
Dolomita	1	3	Ciclohexano	63,9
Dolomita	1	3	Ciclohexano	63,7
Dolomita	1	3	Ciclohexano	63,8
Dolomita	1	3	Ciclohexano	63,7
Dolomita	1	3	Ciclohexano	63,9
Dolomita	1	3	Ciclohexano	63,6
Dolomita	1	3	Ciclohexano	64
Dolomita	1	3	Ciclohexano	64,1
Dolomita	1	3	Ciclohexano	63,8
Dolomita	1	3	Ciclohexano	63,8
Dolomita	1	3	Ciclohexano	63,5
Dolomita	1	3	Ciclohexano	63,7
Dolomita	1	3	N-Hexadecano	103,9
Dolomita	1	3	N-Hexadecano	103,3
Dolomita	1	3	N-Hexadecano	103,8
Dolomita	1	3	N-Hexadecano	101,8
Dolomita	1	3	N-Hexadecano	103,2
Dolomita	1	3	N-Hexadecano	102,8

Dolomita	1	3	N-Hexadecano	102,7
Dolomita	1	3	N-Hexadecano	102,5
Dolomita	1	3	N-Hexadecano	103
Dolomita	1	3	N-Hexadecano	103,7
Dolomita	1	3	N-Hexadecano	102,8
Dolomita	1	3	N-Hexadecano	102,8
Dolomita	1	3	N-Hexadecano	102,7
Dolomita	1	3	N-Hexadecano	102,7
Dolomita	1	3	N-Hexadecano	103,5
Dolomita	1	3	N-Hexadecano	102,2
Dolomita	1	3	N-Hexadecano	101,8
Dolomita	1	3	N-Hexadecano	102
Dolomita	1	3	N-Hexadecano	102,1
Dolomita	1	3	N-Hexadecano	101,2
Dolomita	1	3	Tolueno	52,7
Dolomita	1	3	Tolueno	52,8
Dolomita	1	3	Tolueno	52,8
Dolomita	1	3	Tolueno	52,6
Dolomita	1	3	Tolueno	52,9
Dolomita	1	3	Tolueno	52,8
Dolomita	1	3	Tolueno	52,7
Dolomita	1	3	Tolueno	52,7
Dolomita	1	3	Tolueno	52,8
Dolomita	1	3	Tolueno	52,6
Dolomita	1	3	Tolueno	52,8
Dolomita	1	3	Tolueno	52,7
Dolomita	1	3	Tolueno	52,9
Dolomita	1	3	Tolueno	52,6
Dolomita	1	3	Tolueno	52,8
Dolomita	1	3	Tolueno	52,8
Dolomita	1	3	Tolueno	52,7
Dolomita	1	3	Tolueno	52,7
Dolomita	1	3	Tolueno	52,7
Dolomita	1	3	Tolueno	52,7
Dolomita	1	15	Ciclohexano	93,9
Dolomita	1	15	Ciclohexano	95,6
Dolomita	1	15	Ciclohexano	95,8
Dolomita	1	15	Ciclohexano	95,7
Dolomita	1	15	Ciclohexano	93,7
Dolomita	1	15	Ciclohexano	96,3
Dolomita	1	15	Ciclohexano	95,6
Dolomita	1	15	Ciclohexano	95,7
Dolomita	1	15	Ciclohexano	95,6
Dolomita	1	15	Ciclohexano	95,8
Dolomita	1	15	Ciclohexano	95,8
Dolomita	1	15	Ciclohexano	95,7

Dolomita	1	15	Ciclohexano	95,8
Dolomita	1	15	Ciclohexano	95,8
Dolomita	1	15	Ciclohexano	95,7
Dolomita	1	15	Ciclohexano	95,9
Dolomita	1	15	Ciclohexano	95,9
Dolomita	1	15	Ciclohexano	95,8
Dolomita	1	15	Ciclohexano	95,7
Dolomita	1	15	Ciclohexano	95,8
Dolomita	1	30	Ciclohexano	143
Dolomita	1	30	Ciclohexano	143,2
Dolomita	1	30	Ciclohexano	143,3
Dolomita	1	30	Ciclohexano	143,2
Dolomita	1	30	Ciclohexano	142
Dolomita	1	30	Ciclohexano	141,8
Dolomita	1	30	Ciclohexano	142,8
Dolomita	1	30	Ciclohexano	142,4
Dolomita	1	30	Ciclohexano	142,8
Dolomita	1	30	Ciclohexano	142,5
Dolomita	1	30	Ciclohexano	142,7
Dolomita	1	30	Ciclohexano	142,8
Dolomita	1	30	Ciclohexano	143
Dolomita	1	30	Ciclohexano	143,2
Dolomita	1	30	Ciclohexano	142,4
Dolomita	1	30	Ciclohexano	143,2
Dolomita	1	30	Ciclohexano	143,4
Dolomita	1	30	Ciclohexano	143
Dolomita	1	30	Ciclohexano	142,9
Dolomita	1	30	Ciclohexano	143,4