

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE TECNOLOGIA
PROGRAMA DE ENGENHARIA DE PRODUÇÃO

**MODELAGEM DE DADOS DE FALHAS DE EQUIPAMENTOS DE SUB-
SUPERFICIE EM POÇOS DE PETROLEO DA BACIA POTIGUAR**

Por

MARIA ALDILENE DANTAS

BACHAREL EM ESTATÍSTICA, UFRN, 2004

TESE SUBMETIDA AO PROGRAMA DE ENGENHARIA DE PRODUÇÃO DA
UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE

MESTRE EM CIÊNCIAS EM ENGENHARIA DE PRODUÇÃO

ABRIL, 2008

© 2008 **MARIA ALDILENE DANTAS**
TODOS DIREITOS RESERVADOS.

O autor aqui designado concede ao Programa de Engenharia de Produção da Universidade Federal do Rio Grande do Norte permissão para reproduzir, distribuir, comunicar ao público, em papel ou meio eletrônico, esta obra, no todo ou em parte, nos termos da Lei.

Assinatura do Autor: _____
APROVADO POR:

Profa. Dione Maria Valença, Dra. – Orientador - UFRN, Presidente

Prof. Pledson Guedes de Medeiros, Dr – UFRN, Membro Examinador

Dario José Aloise, Dr – UFRN, Membro Examinador

Profa. Silvia Maria de Freitas, Dra. - UFC, Membro Examinador Externo

Benno Waldemar Assmann, Dr – Engenheiro da PETROBRAS, Membro da Sociedade

Catálogo da Publicação na Fonte. UFRN / SISBI / Biblioteca Setorial Especializada
Especializada do Centro de Ciências Exatas e da Terra – CCET.

Dantas, Maria Aldilene Dantas.

Modelagem de dados de falhas de equipamentos de sub-superfície em poços de
petróleo da Bacia Potiguar / Maria Aldilene Dantas. -- Natal, 2008

117 f. : il.

Orientador: Profa. Dra. Dione Maria Valença.

Dissertação (Mestrado) Universidade Federal do Rio Grande do Norte. Centro de
Tecnologia. Programa de Pós-graduação em Engenharia de Produção.

1. Confiabilidade – Dissertação. 2. Poço de petróleo – Dissertação. 3. Modelo de
regressão Weibull – Dissertação. I. Valença, Dione Maria. II. Título.

RN/UF/BSE-CCET

CDU: 519.248

CURRICULUM VITAE RESUMIDO

Maria Aldilene Dantas é Estatística graduada pela Universidade Federal do Rio Grande do Norte em 2004. Durante a graduação foi bolsista de Iniciação Científica vinculada a Base de Pesquisa: Grupo de Estudos Demográficos – GED – DEST/CCET/UFRN. Participou de vários Congressos Nacionais e Regionais com artigos publicados, destacando os certificados recebidos no 15º SINAPE – Simpósio Nacional de Probabilidade e Estatística, em Águas de Lindóia/SP - 2002, pela publicação do trabalho “Associação da Mortalidade Infantil com Fatores Sociosanitários nas Microrregiões do Rio Grande do Norte”. E pelo trabalho “Distribuição Espacial dos Fatores Socioambientais do Estado do Rio Grande do Norte.” Publicado no XIII Encontro Nacional de Estudos Populacionais, Ouro Preto/MG - 2002. Foi professora Substituta do Departamento de Estatística da UFRN no período de 04/2004 a 12/2005, em ANAIS e CD ROM.

Durante o mestrado foi bolsista do Projeto SAFES – Sistema de Análise de Falhas de Equipamentos de Sub-superfície em poços de petróleo da bacia potiguar. Projeto desenvolvido através de uma parceria entre a Universidade Federal do Rio Grande do Norte – UFRN e a Petrobras. Participou do XXVII ENEGEP – Encontro Nacional de Engenharia de Produção, em Foz de Iguaçu – 2007, com apresentação do trabalho “Elaboração de um plano amostral para medir tempos de falhas em poços de petróleo da PETROBRAS no Rio Grande do Norte e Ceará” e publicação do artigo “Utilização do Controle Estatístico do Processo (CEP) na comparação de velocidades de encapsulamento de medicamento para tuberculose”, Anais CD ROM. Teve trabalho aprovado e apresentado no 18º SINAPE Simpósio Nacional de Probabilidade e Estatística, com o título “O modelo de regressão weibull para estudar o tempo até a primeira falha de equipamento de sub-superfície de poços de petróleo”, 2008, em Estância de São Pedro – SP. Anais e CD ROM.

Submeteu o artigo DANTAS, M. A., FREIRE, M. P. S., VALENÇA, D. M., MEDEIROS, P. G., SILVA, D. N., ALOISE, D. J. *Modelo de Regressão Weibull para estudar dados de falha de equipamentos de sub-superfície em poços petrolíferos* para revista PRODUÇÃO. Aguardando avaliação.

AGRADECIMENTOS

É por dever de justiça que devo expressar profunda gratidão aos coordenadores e colaboradores do projeto Sistema de Análise de Falha de Equipamentos de Sub-superfície – SAFES, do qual este estudo é etapa inicial, através dos professores Dione Maria Valença, Pledson Guedes de Medeiros e Damião Nóbrega da Silva, que durante este período, me orientaram, ensinando a investigar e entender a complexidade de uma pesquisa.

A PETROBRAS pelo fornecimento das informações relevantes para compor este estudo e pelo apoio financeiro.

Ao Programa de Pós-Graduação de Engenharia de Produção da UFRN–PEP, pela oportunidade. Ao coordenador Dario José Aloise pela ajuda e colaboração. A secretária Cleide Batista por sempre está disposta a ajudar.

Aos meus pais, Zé Dantas e Maria Ambrósio, pelo amor e exemplo de vida. A meus irmãos que sempre me apoiaram e colaboraram na minha formação. E a toda a minha família.

A minha orientadora professora Dione Maria Valença e ao meu co-orientador Pledson Guedes de Medeiros pela orientação, paciência e ajuda no desenvolvimento deste trabalho.

Aos colegas do mestrado, pela convivência agradável e a troca de conhecimentos que ajudou na minha formação profissional e pessoal.

A Michel e a Hermes pela grandiosa e fundamental ajuda na análise dos dados e ensinamentos no software R.

E a todos meus amigos presentes e ausentes que contribuíram com energia positiva para a finalização dessa etapa.

Resumo da Tese apresentada a UFRN/PEP como parte dos requisitos necessários para a obtenção do grau de Mestra em Ciências em Engenharia de Produção.

MODELAGEM DE DADOS DE FALHAS DE EQUIPAMENTOS DE SUB-SUPERFÍCIE EM POÇOS DE PETRÓLEO DA BACIA POTIGUAR.

MARIA ALDILENE DANTAS

Abril/2008

Orientador: Profa. Dr. Dione Maria Valença

Co-Orientadora: Prof. Dr. Pledson Guedes de Medeiros

Curso: Mestrado em Ciências em Engenharia de Produção

RESUMO: Esta dissertação apresenta uma aplicação do estudo da confiabilidade em dados de tempo de vida de poços petrolíferos terrestres produtores de óleo da Bacia Potiguar (RN/CE). A variável resposta do nosso estudo é o *tempo* de funcionamento do poço dentro de sua normalidade até apresentar a primeira falha relacionada a equipamentos de subsuperfície, que cause uma parada total no funcionamento do poço. O principal objetivo do estudo foi, com base em um conjunto de dados reais fornecidos pela PETROBRAS, verificar a existência do relacionamento do tempo de vida com diversas características dos poços como, método de elevação utilizado, quantidade de água produzida (BSW Basic Sediments and Water), Razão Gás Óleo (RGO), profundidade de instalação da bomba, unidade operacional de origem, entre outras. Neste trabalho foi feito um estudo retrospectivo com uma amostra composta por 603 poços-colunas que estavam em funcionamento no período de 2000 a 2006. A modelagem probabilística dos dados foi feita através do ajuste do modelo de regressão Weibull. O ajuste deste modelo se mostrou apropriado aos dados e permitiu avaliar, através de testes estatísticos de hipóteses, os fatores significativamente relacionados com o tempo até a falha dos poços.

Palavras-chave: Confiabilidade, poço de petróleo, Modelo regressão Weibull

Abstract of Master Thesis presented to UFRN/PEP as fulfillment of requirements to the degree of Master of Science in Production Engineering

MODELLING OF DATA OF FLAWS OF EQUIPMENTS OF SUB-SURFACE IN OIL WELLS OF THE BACIA POTIGUAR.

MARIA ALDILENE DANTAS

April/2008

Thesis Supervisor: Dione Maria Valença

Thesis Supervisor: Pledson Guedes de Medeiros

Program: Master of Science in Production Engineering

ABSTRACT: This master's thesis presents a reliability study conducted among onshore oil fields in the Potiguar Basin (RN/CE) of Petrobras company, Brazil. The main study objective was to build a regression model to predict the risk of failures that impede production wells to function properly using the information of explanatory variables related to wells such as the elevation method, the amount of water produced in the well (BSW), the ratio gas--oil (RGO), the depth of the production bomb, the operational unit of the oil field, among others. The study was based on a retrospective sample of 603 oil columns from all that were functioning between 2000 and 2006. Statistical hypothesis tests under a Weibull regression model fitted to the failure data allowed the selection of some significant predictors in the set considered to explain the first failure time in the wells.

Key words: reliability, Weibull regression model, onshore oil fields

LISTA DE FIGURAS

FIGURA 1.1 - BOMBEAMENTO ATRAVÉS DE CAVALO DE PAU, NO CAMPO DE SERRA-MACAU,	2
FOTOS:MIRIAM DO NASCIMENTO, JUNHO DE 2002	2
FIGURA 1.2 - BOMBEAMENTO POR CAVIDADES PROGRESSIVAS (BCP).	2
FIGURA 1.3 - Sonda de produção terrestre montada.....	3
FIGURA 2.1 - (A) TODOS OS PACIENTES EXPERIMENTARAM O EVENTO ANTES DO FINAL DO ESTUDO. (B) NO FINAL DO ESTUDO ALGUNS PACIENTES AINDA NÃO HAVIAM EXPERIMENTADO O EVENTO DE INTERESSE.	9
FIGURA 2.2 - FUNÇÕES DE TAXA DE FALHA - - CRESCENTE, — CONSTANTE E - - . DECRESCENTE.....	11
FIGURA 3.1 - COLUNA CONVENCIONAL DE PRODUÇÃO EQUIPADA COM GÁS-LIFT.....	26
FIGURA 3.2 - TIPOS DE COMPLETAÇÃO: (A) SIMPLES, (B) SELETIVA E (C) DUPLA	27
FIGURA 3.3 - SISTEMA DE BOMBEIO MECÂNICO	29
FIGURA 3.4 - SISTEMA DE BOMBEIO POR CAVIDADE PROGRESSIVA-BCP	30
FIGURA 4.1 - DISTRIBUIÇÃO DAS MEDIDAS DE PROFUNDIDADE DA BOMBA SEGUNDO A UNIDADE ADMINISTRATIVA E O MÉTODO DE ELEVAÇÃO.....	41
FIGURA 5.1 - NÚMERO DE POÇOS-COLUNA AMOSTRADOS POR ESTADO DE LOCALIZAÇÃO.....	49
FIGURA 5.2 - NÚMERO DE POÇOS-COLUNA AMOSTRADOS POR MÉTODOS DE ELEVAÇÃO, NO PERÍODO DE 2000 A 2006	49
FIGURA 5.3 - NÚMERO DE POÇOS-COLUNA AMOSTRADOS POR ATIVO DE PRODUÇÃO.....	50
FIGURA 5.4 - PRODUÇÃO BASE POR MÉTODO DE ELEVAÇÃO	52
FIGURA 5.5 - PRODUÇÃO BASE POR UNIDADE OPERACIONAL	53
FIGURA 5.6 - BSW POR MÉTODO DE ELEVAÇÃO	54
FIGURA 5.7 - BSW POR UNIDADE OPERACIONAL.....	55
FIGURA 5.8 - RGO POR MÉTODO DE ELEVAÇÃO	56
FIGURA 5.9 - RGO POR UNIDADE ADMINISTRATIVA	57
FIGURA 5.10 - PROFUNDIDADE DO POÇO POR MÉTODO DE ELEVAÇÃO.....	57
FIGURA 5.11 - PROFUNDIDADE DO POÇO POR UNIDADE ADMINISTRATIVA.....	58
FIGURA 5.12 - PROFUNDIDADE DA BOMBA POR MÉTODO DE ELEVAÇÃO.....	59
FIGURA 5.13 - PROFUNDIDADE DA BOMBA POR UNIDADE ADMINISTRATIVA	60
60	
FIGURA 5.14 - HISTOGRAMA DO TEMPO ATÉ A PRIMEIRA FALHA DO POÇO	62
FIGURA 5.15 - CURVA DE SOBREVIVÊNCIA ESTIMADA A PARTIR DO ESTIMADOR DE KAPLAN-MEIER.....	63
63	
FIGURA 5.18 - CURVA DE SOBREVIVÊNCIA ESTIMADA PELO MÉTODO KAPLAN-MEIER, POR FAIXA DE IDADE DO POÇO (POÇO NOVO < 6.72 ANOS, POÇO VELHO ≥6.72 ANOS).....	65
FIGURA 5.19 - CURVA DE SOBREVIVÊNCIA ESTIMADA PELO MÉTODO KAPLAN-MEIER, POR NÍVEL DE RGO (BAIXO < 4 m ³ , ALTO ≥ 4 m ³).....	66
FIGURA 5.20 - CURVA DE SOBREVIVÊNCIA ESTIMADA PELO MÉTODO KAPLAN-MEIER, POR MÉTODO DE ELEVAÇÃO (BM E BCP)	66
FIGURA 5.21 - CURVA DE SOBREVIVÊNCIA ESTIMADA PELO MÉTODO KAPLAN-MEIER, POR UNIDADE OPERACIONAL	67
FIGURA 5.22 - CURVA DE SOBREVIVÊNCIA ESTIMADA PELO MÉTODO DE KAPLAN-MEIER, POR NÍVEL DE PROFUNDIDADE DA BOMBA (BAIXA < 490M, ALTA ≥ 490M).....	68
FIGURA 5.23 - CURVA DE SOBREVIVÊNCIA ESTIMADA PELO MÉTODO DE KAPLAN-MEIER, POR NÍVEL DE PROFUNDIDADE DO POÇO (BAIXA < 675M, ALTA ≥ 675M).....	69
FIGURA 5.24 - CURVA DE SOBREVIVÊNCIA ESTIMADA PELO MÉTODO DE KAPLAN-MEIER, PARA CADA UNIDADE OPERACIONAL , POR NÍVEL PRODUÇÃO DE ÓLEO (BAIXA < 2.20m ³ /DIA, ALTA ≥ 2.20m ³ /DIA).	69
FIGURA 5.25 - CURVAS DE SOBREVIVÊNCIA ESTIMADA PELO MÉTODO DE KAPLAN-MEIER PARA CADA UNIDADE OPERACIONAL, POR NÍVEL DE BSW (BAIXO < 79,40%, ALTO ≥79,40%)	70
FIGURA 5.26 - CURVAS DE SOBREVIVÊNCIAS ESTIMADA PELO MÉTODO DE KAPLAN-MEIER PARA CADA UNIDADE OPERACIONAL, POR MÉTODO DE ELEVAÇÃO.	71
FIGURA 5.27 - CURVAS DE SOBREVIVÊNCIA ESTIMADA PELO MÉTODO DE KAPLAN-MEIER PARA CADA UNIDADE OPERACIONAL, EM DIFERENTES NÍVEIS POR NÍVEL DE RGO (BAIXO < 4 m ³ DIA ALTO ≥ 4 m ³ DIA)	71
FIGURA 5.28 - CURVAS DE SOBREVIVÊNCIA ESTIMADA PELO MÉTODO DE KAPLAN-MEIER PARA CADA MÉTODO DE ELEVAÇÃO, POR NÍVEL DE RGO (BAIXO < 4 m ³ DIA ALTO ≥ 4 m ³ DIA).....	72
FIGURA 5.29 - CURVAS DE SOBREVIVÊNCIAS ESTIMADA PELO MÉTODO DE KAPLAN-MEIER PARA CADA MÉTODO DE ELEVAÇÃO, POR NÍVEL DE PROFUNDIDADE DO POÇO (BAIXO < 675 M, ALTO ≥ 675 M).....	73
FIGURA 5.30 - GRÁFICO DE ADEQUAÇÃO DO MODELO WEIBULL PARA OS DADOS DE TEMPO ATÉ A PRIMEIRA FALHA DOS POÇOS-COLUNAS DA BACIA POTIGUAR, NO PERÍODO DE 2000 A 2006.....	74

FIGURA 5.31 - CURVA DE SOBREVIVÊNCIA ESTIMADA BASEADA NO MODELO WEIBULL.....	75
FIGURA 5.32 - CURVA DE SOBREVIVÊNCIA ESTIMADA BASEADA NO KM E NO MODELO WEIBULL	76
FIGURA 5.33 - CURVA DE SOBREVIVÊNCIA DOS RESÍDUOS OBTIDA POR KAPLAN-MEIER E PELO MÉTODO EXPONENCIAL	78
FIGURA 5.34 - GRÁFICO DE PARES DE PONTOS DOS RESÍDUOS OBTIDOS POR KAPLAN-MEIER E PELO MÉTODO EXPONENCIAL	78
FIGURA 5.35 - CURVAS DE SOBREVIVÊNCIA ESTIMADA ATRAVÉS DO MODELO DE REGRESSÃO WEIBULL POR UNIDADE OPERACIONAL EM DIFERENTES NÍVEIS DE BSW (BSW =50 E BSW = 80).....	79
FIGURA 5.36 - CURVAS DE SOBREVIVÊNCIA ESTIMADA ATRAVÉS DO MODELO DE REGRESSÃO WEIBULL PARA POÇO COM MÉTODO DE ELEVAÇÃO BM E BCP POR UNIDADE OPERACIONAL	80

LISTA DE TABELAS

TABELA 2.1 - ILUSTRAÇÃO DOS PROCEDIMENTOS CITADOS	24
TABELA 4.1 - DEFINIÇÃO E TAMANHO DOS ESTRATOS PARA A POPULAÇÃO DE 4.358 COLUNAS DE PRODUÇÃO.	41
TABELA 4.2 - ALOCAÇÃO DA AMOSTRA E PESOS AMOSTRAIS	42
TABELA 4.3 - NÚMERO DE COLUNAS AMOSTRADAS, POR UNIDADE ADMINISTRATIVA.	42
TABELA 4.4 - NÚMERO DE COLUNAS AMOSTRADAS, POR MÉTODO DE ELEVAÇÃO	43
TABELA 4.5 - NÚMERO DE COLUNAS AMOSTRADAS, POR UNIDADE ADMINISTRATIVA E MÉTODO DE ELEVAÇÃO	43
TABELA 5.17 - NÚMERO DE FALHAS E CENSURAS POR MÉTODO DE ELEVAÇÃO E UNIDADE OPERACIONAL OCORRIDAS NOS POÇOS-COLUNAS AMOSTRADOS, NO PERÍODO DE 2000 A 2006.....	60
TABELA 5.18 - MEDIDAS DESCRITIVAS PARA O TEMPO ATÉ A FALHA DO POÇO-COLUNA CORRESPONDENTE A AMOSTRA ANALISADA NO PERÍODO DE 2000 A 2006.....	61
TABELA 5.19 - ESTIMATIVA DAS MEDIDAS DESCRITIVAS PARA O TEMPO ATÉ A FALHA DO POÇO-COLUNA, POR COVARIÁVEL, NO PERÍODO DE 2000 A 2006.....	61

SUMÁRIO

1. INTRODUÇÃO.....	1
1.1 - RELEVÂNCIA DO ESTUDO.....	5
1.2 - TIPO DE PESQUISA	5
1.3 - ESTRUTURA DO TRABALHO.....	6
2. ANÁLISE DE DADOS DE TEMPO DE VIDA.....	7
2.1 - CONCEITOS BÁSICOS DE ANÁLISE DE SOBREVIVÊNCIA.....	7
2.2 - CARACTERÍSTICAS DAS DISTRIBUIÇÕES DOS DADOS	9
2.3 - ESTIMAÇÃO NÃO-PARAMÉTRICA DA FUNÇÃO DE SOBREVIVÊNCIA	11
2.3.1 - FUNÇÃO DE SOBREVIVÊNCIA EMPÍRICA	12
2.3.2 - ESTIMADOR KAPLAN-MEIER	12
2.4 - MODELOS PARAMÉTRICOS.....	13
2.4.1 - DISTRIBUIÇÃO EXPONENCIAL	13
2.4.2 - DISTRIBUIÇÃO DE WEIBULL.....	14
2.5 - INFERÊNCIA EM MODELOS PARAMÉTRICOS.....	16
2.5.1 - O MÉTODO DE MÁXIMA VEROSSIMILHANÇA	16
2.6 - AJUSTE DOS DADOS A UMA DISTRIBUIÇÃO EXPONENCIAL E WEIBULL	19
2.6.2 - DISTRIBUIÇÃO WEIBULL	20
2.7 - MODELO DE REGRESSÃO WEIBULL	21
2.8 - ESTRATÉGIA PARA A SELEÇÃO DE COVARIÁVEIS NO MODELO	23
3. CONCEITOS FUNDAMENTAIS DE POÇOS DE PETRÓLEO	25
3.1 - CONCEITOS PRELIMINARES	25
3.2 - BOMBEIO MECÂNICO (BM)	28
3.3 - BOMBEIO POR CAVIDADE PROGRESSIVA (BCP)	29
3.4 - ACOMPANHAMENTO DA PRODUÇÃO	30
3.5 - OPERAÇÕES DE MANUTENÇÃO EM POÇOS.....	32
3.6 - DEFINIÇÃO DO ESTADO DOS POÇOS	33
3.7 - CLASSIFICAÇÃO DOS POÇOS	35
3.8 - SISTEMAS DE INFORMAÇÕES DA PETROBRAS	36
4. PLANO AMOSTRAL PARA SELEÇÃO DOS POÇOS.....	38
4.1 - COLETA DE DADOS	38
4.2 - POPULAÇÃO CONSIDERADA NO ESTUDO	39
4.3 - AMOSTRA.....	39
4.3.1 - TAMANHO E ALOCAÇÃO DA AMOSTRA	42
4.4 - SELEÇÃO DA AMOSTRA.....	43
5. RESULTADOS	44
5.1 - ANÁLISE EXPLORATÓRIA DOS DADOS.....	44
5.2 - ANÁLISE NÃO-PARAMÉTRICA	62
5.2.1 - ESTIMADOR KAPLAN-MEIER	63
5.2.2 - COMPARAÇÃO ENTRE AS CURVAS DE SOBREVIVÊNCIA	63
5.3 - AJUSTE DO MODELO WEIBULL E DETERMINAÇÃO DOS PARÂMETROS	73
5.3.1 - AJUSTE DO MODELO DE REGRESSÃO WEIBULL.....	76
6. CONCLUSÃO E RECOMENDAÇÕES	82
6.1 - PESQUISA BIBLIOGRÁFICA.....	82
6.2 - METODOLOGIA DA PESQUISA.....	83
6.3 - RESULTADO DA PESQUISA	84
6.4 - LIMITAÇÕES DO TRABALHO	85
6.5 - DIREÇÕES DA PESQUISA	85
6.6 - CONCLUSÃO.....	85
REFERÊNCIAS BIBLIOGRÁFICAS.....	88

1. INTRODUÇÃO

Petróleo na Bacia Potiguar - Breve Histórico

Após a criação da Petrobras, em 1953, deu-se início às pesquisas de petróleo no Brasil. Embora no Rio Grande do Norte já houvesse registros informando possível existência de petróleo no território, as primeiras perfurações só ocorreram em 1956 e as expectativas com relação à produção dos poços não foram satisfatórias. Somente em 1956, após um estudo mais aprofundado, a Petrobras informou que a Bacia Potiguar tinha condições de produzir petróleo. A primeira perfuração com produção lucrativa ocorreu somente em 1977. Em seguida ocorreram várias descobertas de poços e assim, em 1981, a Bacia estava formada com quatro grandes campos produzindo diariamente 157 barris de petróleo (Revista Petrobras, 2005)

Em 1985, a Petrobras descobriu na região entre Mossoró e Areia Branca um novo campo de petróleo, entrando em produção em de janeiro de 1986. Era o campo de Canto do Amaro, atualmente o maior campo terrestre produtor do Brasil. Com a nova descoberta a produção saltou dos 59.000 barris/dia, em 1985, para 71.200 barris/dia, ao final de 1986.

Ultimamente, a Bacia Potiguar produz aproximadamente 78.000 barris/dia, provenientes de 60 campos petrolíferos com número de poços variável, situados no RN e seis campos produtores no CE, abrangendo aproximadamente 4.200 poços nos dois Estados e correspondendo a 4 % da produção nacional. Estes números colocam a Bacia Potiguar não somente em primeiro lugar no Brasil em termos de produção terrestre, mas também na segunda colocação na produção de petróleo.

A elevação dos fluidos na maioria desses poços é feita de forma artificial, pois cerca de 98% dos poços petrolíferos são não surgentes, ou seja, não possuem pressão suficiente para que os fluidos atinjam a superfície. A elevação pode ser realizada através de diversas técnicas. Na Bacia Potiguar terrestre as técnicas mais usadas são o bombeio mecânico (BM) ou “Cavalo de Pau” (Figura 1.1) e o bombeio por cavidades progressivas (BCP) (Figura 1.2). Essas técnicas exigem a instalação de equipamentos que trabalharão por um

longo período junto ao poço, realizando o bombeamento dos fluidos, podendo atingir uma idade produtiva de 15 anos ou mais. Ao longo do tempo, poços vão falhando e, como consequência, surge a necessidade de serviços de manutenção (limpeza, restauração, estimulação, avaliação) e completação (término para exploração). Estes serviços são realizados por Sondas de Produção Terrestres (SPT) móveis, as quais são terceirizadas a um custo muito elevado, por esta razão, há um número muito pequeno de sondas atuando em toda a área (Figura 1.3).

Figura 1.1 - Bombeamento através de cavalo de pau, no campo de Serra-Macau,



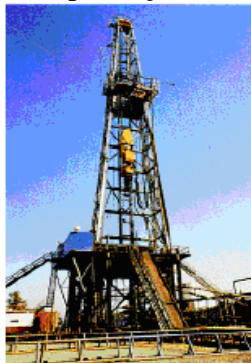
Fotos: Miriam do Nascimento, Junho de 2002.

Figura 1.2 - Bombeamento por cavidades progressivas (BCP).



Fotos: Miriam do Nascimento, julho de 2002.

Figura 1.3 - Sonda de produção terrestre montada



Fonte: Thomas, 2001

As características dos poços serão chamadas de variáveis, ou mais especificamente, *covariáveis*, e avaliamos a associação de diversas características (covariáveis) em relação ao tempo até a primeira falha do poço.

Revisão Bibliográfica

O presente estudo está contextualizado na área de confiabilidade, especificamente na análise de dados de equipamentos de sub-superfície de poços de petróleo da Petrobras, nos estados do Rio Grande do Norte e Ceará (UN-RNCE - Unidade de Negócio de Exploração e Produção do Rio Grande do Norte e Ceará).

A análise de tempo de falha é um procedimento estatístico utilizado para analisar dados que correspondem ao tempo de ocorrência de um evento de interesse, tal como o tempo até a falha de um equipamento (Kleibaum, 1996). E, segundo Allisson (1995), um evento é conceituado como sendo uma mudança qualitativa situada no tempo. Neste estudo, o evento de interesse corresponde ao tempo até a primeira falha do poço.

Na área médica, esta técnica é conhecida como Análise de Sobrevivência. Segundo Colosimo e Giolo (2006) é uma técnica de ampla utilização e é uma das áreas da estatística que mais cresceu nos últimos anos. Na literatura existem muitas técnicas apropriadas para análise de tempo de falha (ver por exemplo, Lawles, 1982; Nelson, 1990).

Segundo Lawless (1982) a maioria dos trabalhos estatísticos na área de confiabilidade foi aplicado ao tempo de vida de equipamentos e sistemas.

Lindqvist *et al* (1988) aplicou o modelo de riscos proporcionais de Cox (1972) a dados de falhas das principais válvulas de segurança de sub-superfície utilizadas em 26 campos de óleo e gás no Mar do Norte.

Accioly e Martins (1999) realizaram análises das intervenções no campo de Marlim e Albacora, na Bacia de Campos, Brasil. Eles aplicaram a distribuição Weibull aos dados de tempo entre as intervenções e ao tempo de produção dos poços que não sofreram intervenção.

Esta mesma distribuição foi utilizada por Bardy (2000) em uma análise estatística dos dados de tempos de vida (tempos de operação) de um conjunto de bombeio centrífugo submerso (BCS), instalados nos poços do Pólo Nordeste da Bacia de Campos.

César (2001) aplicou modelo de regressão de Cox (1992) e os modelos paramétricos Weibull e exponencial para analisar dados de tempo de falhas dos poços da região do Alto Rodrigues pertencentes à Bacia Potiguar Terrestre.

Frota (2003) utilizou técnicas de confiabilidade para analisar dados de um período de 12 anos de intervenções em poços marítimos, da Bacia de Campos, Brasil, com a finalidade de identificar as causas das intervenções, modelar falhas e analisar o tempo de falha. A distribuição de probabilidade utilizada na análise estatística dos dados foi a distribuição Weibull.

A revisão bibliográfica mostrou poucas publicações relacionadas à análise de tempo de falha de poços, considerando-o como um sistema. No entanto, há várias pesquisas referentes a análise de falha de equipamento específico do poço utilizando-se de técnicas de análise de sobrevivência. Exemplo de equipamentos comuns a este tipo de estudo: falha em bombeio mecânico, falha em bomba Centrifugo, válvulas de segurança de sub-superfície, entre outros.

Objetivo

O objetivo deste trabalho é fazer um estudo retrospectivo sobre informações relativas a poços terrestres produtores de óleo da Bacia potiguar, no período de 2000 a 2006, cujo método de elevação é BM e BCP, para ajustar um modelo que indique o risco de falha, de acordo com as características dos poços. Nesta análise serão consideradas as variáveis relevantes para falhas e algumas características das falhas. Este trabalho está associado ao

projeto SAFES (Sistema de Análise de Falhas de Equipamentos de Sub-superfície) que é desenvolvido através de uma parceria entre a Universidade Federal do Rio Grande do Norte – UFRN e a Petrobras.

1.1 - RELEVÂNCIA DO ESTUDO

1.1.1 Relevância Teórica

Científica: Este estudo visa dar subsídios para a solução de problemas de manutenção de poços de petróleo, através da construção de um modelo que permita a modelagem do tempo de falha e determinação das variáveis influentes.

1.1.2 Relevância Prática

Devido a grande quantidade de poços existentes na Bacia Potiguar é de grande relevância conhecer o risco de falha para definir quais poços deve ter uma análise detalhada de falha. Uma análise preventiva das ocorrências de falhas dos poços possibilitará um acompanhamento melhor dos poços que apresentam maiores risco de falha e também garantirá que materiais requeridos para o restabelecimento da produção estejam disponíveis no momento da falha do poço, reduzindo assim o tempo de manutenção do poço.

O estudo aborda uma amostra de poços, escolhido de forma a representar todos os poços da bacia RN/CE, possibilitando, desta forma, que o resultado seja utilizado para os demais poços desta bacia.

1.2 - TIPO DE PESQUISA

Neste estudo será tratada a *pesquisa exploratória e a explicativa*. Segundo Gil (1999) a pesquisa exploratória tem como principal finalidade proporcionar uma visão geral, de tipo aproximativo, acerca de determinado fato. É usada quando se tem o primeiro contato com os dados, a finalidade é obter uma nova percepção do tema em estudo. Ainda segundo o autor, a pesquisa se caracteriza como explicativa por ter como preocupação central,

identificar os fatores que determinam ou que contribuem para a ocorrência dos fenômenos. O modelo desenvolvido neste estudo será utilizado para auxiliar na determinação dos fatores que influenciam o tempo até a ocorrência da primeira falha a um poço-coluna produtor de óleo da bacia potiguar.

O método utilizado será do tipo quantitativo. Segundo Richardson *et. al* (1985 p. 29), no estudo quantitativo há uma preocupação com a medição objetiva e a quantificação dos resultados, evitando distorções na etapa de análise e interpretação dos dados, garantindo margem de segurança em relação às inferências obtidas.

1.3 - ESTRUTURA DO TRABALHO

Esta dissertação está estruturada em seis capítulos. O primeiro, já descrito acima, fornece uma visão geral do contexto em que o trabalho está inserido, apresentando os objetivos e a contribuição do estudo.

O Capítulo 2 apresenta a fundamentação teórica preliminar ao estudo de confiabilidade (análise de sobrevivência). As distribuições de probabilidades na análise estatística de dados de sobrevivência e o modelo de regressão paramétrico considerado para ajustar o conjunto de dados em questão.

O Capítulo 3 apresenta alguns conceitos, tais como: definição do petróleo, os equipamentos que compõem o poço, o método de elevação de interesse nesse estudo, os estados o qual um poço pode se encontrar durante o período de produção, as intervenções realizadas nos poços, a classificação dos poços e as variáveis envolvidas no estudo.

No Capítulo 4 tem-se, a forma de coleta e a metodologia do plano amostral utilizado para selecionar os poços-coluna deste estudo.

O Capítulo 5 mostra os resultados da aplicação do modelo paramétrico utilizado na análise de dados de falhas de equipamentos de sub-superfície.

O Capítulo 6 apresenta as conclusões da aplicação da análise e recomendações para trabalhos futuros.

2. ANÁLISE DE DADOS DE TEMPO DE VIDA

O presente estudo, que diz respeito à análise das ocorrências de falhas dos poços, pode ser abordado como uma aplicação da área da Estatística que trata de *dados de tempo de vida*. O interesse neste trabalho é modelar estatisticamente o tempo até a ocorrência da primeira falha do poço e de avaliar a influência de diversos fatores na permanência destes em funcionamento. Este capítulo está organizado em 8 seções e apresenta conceitos de Análise de Sobrevivência e de Confiabilidade que se referem às áreas da Estatística que tratam da análise de *Dados de Tempo de Vida*. Inicialmente (seções 2.1 e 2.2) são apresentados alguns conceitos básicos como confiabilidade, tempo de falha, censura e tipo de censura, assim como características probabilísticas da variável aleatória de interesse (função de densidade, função de sobrevivência e função risco). A Seção 2.3 apresenta métodos não paramétricos básicos para estimar a função de sobrevivência. As Seções seguintes descrevem o modelo paramétrico Weibull, bastante usado em aplicações de dados das engenharias, e procedimento para estimação e teste de parâmetros deste modelo, incluindo a situação em que a população de interesse é heterogênea (presença de covariáveis).

2.1 - CONCEITOS BÁSICOS DE ANÁLISE DE SOBREVIVÊNCIA

Atualmente, estudos envolvendo as técnicas de análise de sobrevivência vêm sendo apresentados com grande frequência. Esta técnica é empregada quando se pretende analisar dados relacionados ao tempo até a ocorrência de algum evento de interesse. Este tempo é o intervalo decorrido entre o tempo inicial do estudo, ou início que o indivíduo entrou no estudo, até a ocorrência do evento desejável, conhecido como **falha**. Este tempo é denominado **tempo até a falha ou tempo de vida**. Neste estudo, o tempo inicial foi contado a partir de 01/01/2000. O tempo de operação foi medido em horas, a partir do tempo inicial até a primeira falha do poço-coluna.

De acordo com Colosimo (2001) a denominação “análise de sobrevivência” originou-se na medicina, uma vez que o foco da maioria dos trabalhos é comparar o tempo médio de vida de determinado grupo de pacientes, ou seja, o tempo de sobrevivência. Por outro lado, na engenharia, onde se estuda o tempo de vida de componentes de um equipamento, convencionou-se a denominação de “confiabilidade”.

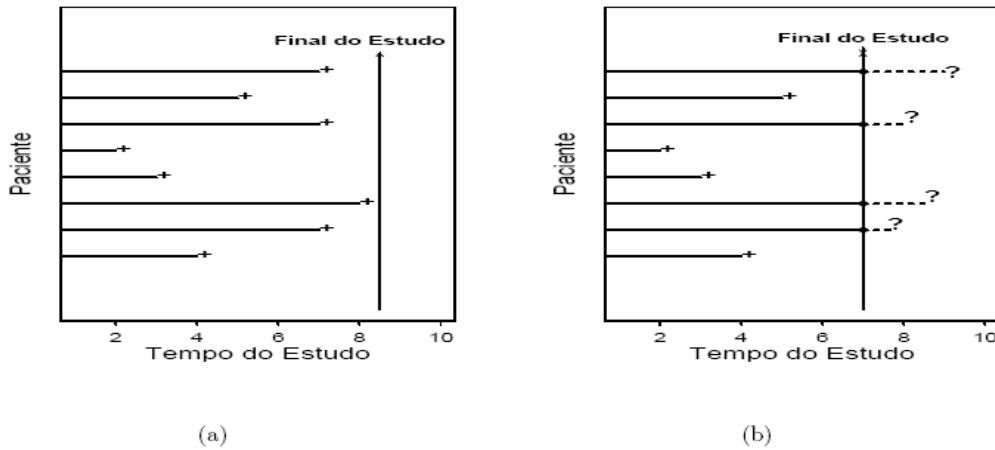
Segundo Accioly (1995) confiabilidade de um equipamento ou de um produto industrial é a probabilidade de que ele execute suas funções com sucesso, sob um período específico de tempo, sem apresentar parada para manutenção ou falha. Um estudo de confiabilidade consiste basicamente em ajustar modelos estatísticos que permitam avaliar o risco de falha do equipamento.

A variável de interesse no estudo de confiabilidade, o tempo de sobrevivência ou até a falha, é uma variável contínua e positiva. A principal característica de um conjunto de dados de confiabilidade é a presença de dois componentes: falhas e censuras, os quais constituem a resposta:

- ✚ Falha – é a ocorrência de algum evento de interesse. Em estudos da área médica, o evento pode ser a morte do paciente ou a recidiva da doença. Nesta dissertação, a falha corresponde a parada de funcionamento do poço, devido alguma anormalidade dos equipamentos de subsuperfície.
- ✚ Censuras - A principal característica de dados de sobrevivência é a presença de censuras que ocorrem quando no final do ensaio nem todas as unidades falharam ou quando há perda de informação decorrente de não ter se observado o tempo do indivíduo que pertence ao grupo em estudo.

A Figura 2.1, a seguir, ilustra a ocorrência de falha e censura em um estudo hipotético com tempo de estudo fixado previamente.

Figura 2.1 - (a) Todos os pacientes experimentaram o evento antes do final do estudo. (b) No final do estudo alguns pacientes ainda não haviam experimentado o evento de interesse.



Segundo Colosimo e Giolo (2006) os três tipos de censuras mais comuns são:

Censura Tipo I – O tempo de estudo é pré-determinado de tal maneira que o tempo de vida só será conhecido se for menor que o tempo determinado, (Figura 2.1-b).

Censura tipo II – O tipo II é aquele onde n indivíduos são colocados em teste e o estudo só termina após ter ocorrido o evento de interesse em um número r ($r \leq n$) pré-fixado de indivíduos.

Censura aleatória - Ocorre por exemplo, num experimento médico, quando os pacientes entram no estudo de forma aleatória de acordo com a data do diagnóstico e, no final do experimento, as observações são censuradas por motivos aleatórios que poderiam ser, entre outros, a morte do paciente por uma causa diferente da estudada ou a perda de contato com o mesmo.

2.2 - CARACTERÍSTICAS DAS DISTRIBUIÇÕES DOS DADOS

Seja T uma variável aleatória (v.a) positiva e contínua que representa o tempo de vida de um equipamento, sendo $T > 0$. T denota a variável aleatória associada ao tempo de sobrevivência e t um valor específico assumido pela variável aleatória T .

Para expressar o comportamento dessa variável recorre-se a três funções importantes que são: a função de densidade de probabilidade $f(t)$; a função de sobrevivência $S(t)$ e a função de risco $h(t)$.

- Função Densidade de Probabilidade

Representa aproximadamente a probabilidade de o indivíduo sobreviver no intervalo $[t, t + \Delta t)$ por unidade de tempo, que pode ser expressa por:

$$f(t) = \lim_{\Delta t \rightarrow 0^+} \frac{P(t \leq T < t + \Delta t)}{\Delta t} = \frac{dF(t)}{dt}. \quad (2.1)$$

A função de distribuição acumulada de T é definida por:

$$F(t) = P(T \leq t) = \int_0^t f(u) du, \quad (2.2)$$

que representa a probabilidade do indivíduo sobreviver até algum tempo t qualquer. Tem-se ainda que:

$$f(t) \geq 0 \quad \text{e} \quad \int_0^{\infty} f(t) dt = 1$$

- Função de Sobrevivência

A função de sobrevivência, denotada por $S(t)$, representa a probabilidade de um indivíduo sobreviver pelo menos até certo tempo t , sendo representado por

$$S(t) = P(T \geq t) = \int_t^{\infty} f(x) dx = 1 - F(t) \quad (2.3)$$

Esta é uma função monótona decrescente com as seguintes propriedades:

$$S(0) = 1 \quad \text{e} \quad \lim_{t \rightarrow \infty} S(t) = 0$$

- Função de Risco

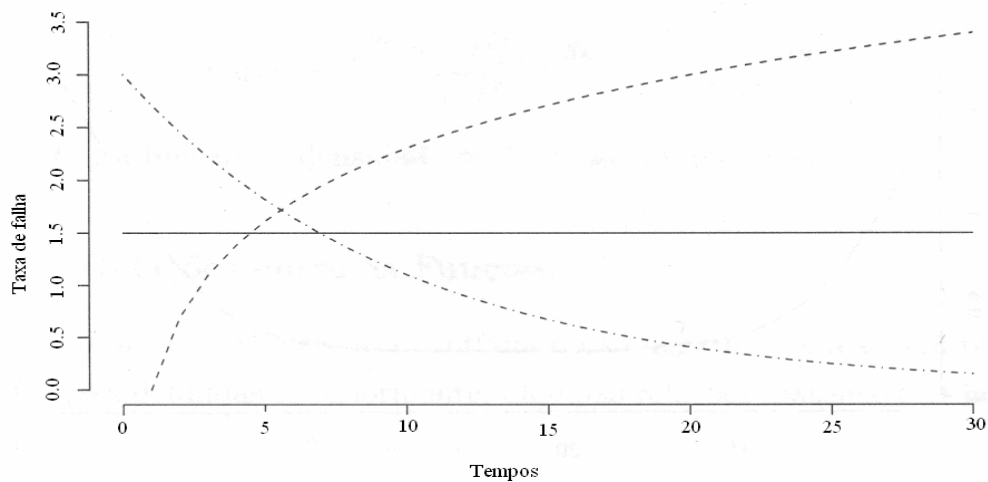
A função de risco, denotada por $h(t)$, especifica a taxa de falha instantânea no tempo t , sendo definida com

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[(t \leq T < t + \Delta t) | T \geq t]}{\Delta t} = \frac{f(t)}{S(t)}, \quad (2.4)$$

Uma distribuição do tempo de falha pode ser compreendida através da função risco, também chamada de *taxa de falha*, pois ela reflete o efeito do tempo sobre os indivíduos que não falharam.

A Figura 2.2 mostra três representações para uma função de risco. A função crescente indica que a taxa de falha aumenta com o transcorrer do tempo. A função constante indica que a taxa não se altera com o passar do tempo. A função decrescente mostra que a taxa de falha diminui à medida que o tempo passa.

Figura 2.2 - Funções de taxa de falha - - crescente, — constante e -.- decrescente.



2.3 - ESTIMAÇÃO NÃO-PARAMÉTRICA DA FUNÇÃO DE SOBREVIVÊNCIA

Em análise de sobrevivência é de interesse estimar a função de sobrevivência da variável em análise. Existem técnicas paramétricas e não paramétricas para se estimar esta função. Segundo Colosimo e Giolo (2006) as técnicas não paramétricas são utilizadas quando não supomos nenhum modelo teórico para o tempo de falha. Nesta seção, serão apresentados os dois métodos de estimação não-paramétricos, a função de sobrevivência empírica e o estimador Kaplan-Meier, eles funcionam como medidas descritivas do conjunto de dados.

2.3.1 - FUNÇÃO DE SOBREVIVÊNCIA EMPÍRICA

Na ausência de censuras, um estimador não paramétrico para a função de sobrevivência $S(t)$ é obtido através da Função de Sobrevivência Empírica, $\tilde{S}(t)$, que é definida por:

$$\tilde{S}(t) = \frac{n^\circ \text{ de observações que não falharam até o tempo } t}{n^\circ \text{ de observações no estudo}}. \quad (2.5)$$

2.3.2 - ESTIMADOR KAPLAN-MEIER

Na prática, grande parte dos dados de sobrevivência apresenta observações censuradas o que indica a necessidade de técnicas especializadas para incorporar as informações contidas nestas observações. Na aplicação deste trabalho, a censura ocorre quando no fim do estudo o poço não apresentou falha relacionada a equipamentos de subsuperfície ou quando o poço foi desativado antes de apresentar falha. Estes dados são chamados de censuras e devem ser inclusos na análise estatística.

O método mais conhecido e utilizado de estimação não paramétrica de $S(t)$, com censuras presentes na amostra, é o método de Kaplan-Meier ou estimador produto-limite, proposto por Kaplan-Meier (1958), sendo definido da seguinte forma: considerando n indivíduos e $r \leq n$ tempo distintos $t_1 < t_2 < \dots < t_r$, onde t_j o tempo onde ocorrem falhas, d_j o número de falhas em t_j e n_j os números de observações sob risco (não falhou e não foi censurado) até o tempo t_j (exclusive), isto é, n_j é o número de itens em risco no instante imediatamente anterior a t_j . Conforme dados em Lawless (1982) e citado por Freire (2006), o \hat{S}_{KM} é dado por:

$$\tilde{S}_{KM}(t) = \prod_{j=1}^k \left(\frac{n_j - d_j}{n_j} \right), \text{ para } t_{(k)} \leq t < t_{(k+1)}, k = 1, 2, 3, \dots, r. e, \quad (2.6)$$

$$\hat{S}_{KM}(t) = 1, \text{ se } t < t_{(1)} \quad \hat{S}_{KM}(t) = 0, \text{ se } t > t_{(r)}$$

2.4 - MODELOS PARAMÉTRICOS

Existem diversos modelos paramétricos que têm se mostrado bastante adequados para descrever os tempos de vida de produtos/equipamentos, ou tempos de vida de pacientes. A literatura de análise de sobrevivência apresenta vários modelos paramétricos que tratam de dados de vida (ver Lawless, (1982), Bain (1978)). Estes modelos são chamados “*modelos paramétricos de probabilidade para o tempo de falha*”. Neste trabalho damos destaque apenas ao modelo Weibull e ao modelo exponencial, que representa um caso particular deste. Este modelo é considerado por vários autores o mais utilizado nas aplicações em estudos de confiabilidade.

2.4.1 - DISTRIBUIÇÃO EXPONENCIAL

A distribuição exponencial com um parâmetro, por ter uma estrutura simples, é muito utilizada. É a única distribuição que se caracteriza por ter uma função de taxa de falha/ou risco constante.

Se T tem distribuição exponencial com parâmetro $\alpha > 0$, sua função densidade de probabilidade é dada por:

$$f(t) = \frac{1}{\alpha} \exp\left\{-\left(\frac{t}{\alpha}\right)\right\}, \quad t \geq 0, \quad (2.7)$$

em que o parâmetro $\alpha > 0$ é o tempo médio de vida. O parâmetro α tem a mesma unidade do tempo de falha t (se t é medido em horas, α também será fornecido em horas).

A função de sobrevivência $S(t)$ e a função de risco $h(t)$ são dadas respectivamente, por:

$$S(t) = \exp\left\{-\left(\frac{t}{\alpha}\right)\right\} \quad \text{e para } t > 0. \quad (2.8)$$

$$\text{e} \quad h(t) = \frac{1}{\alpha} \quad \text{para } t > 0; \quad (2.9)$$

A função risco, $h(t)$, é constante, sendo esta, a principal característica desta distribuição. Isto significa que, tanto uma unidade velha quanto uma nova, que ainda não

falharam, têm o mesmo risco de falhar em um intervalo futuro (Colosimo e Giolo, 2006). Isso relaciona-se com a falta de memória da distribuição exponencial.

Outras características de interesse são: a média, a variância e os percentis. A média e variância de T são respectivamente α e α^2 . O percentil 100p% corresponde ao tempo em que 100p% dos produtos ou indivíduos falharam. Por exemplo, para $p = 0.7$, o percentil $100 \times 0.7 = 70\%$, representa o tempo em que 70% dos indivíduos falharam. Este valor, denotado por t_p , satisfaz $P(T \leq t_p) = p$. Assim se $T \sim \exp(\alpha)$ tem-se $t_p = \alpha \log(1 - p)$.

2.4.2 - DISTRIBUIÇÃO WEIBULL

A distribuição Weibull, proposta por Weibull (1954), representa uma generalização da distribuição exponencial. Segundo Colosimo e Giolo (2006), este modelo vem sendo usado com bastante frequência, entres outras aplicações, em estudo nas áreas médicas e industriais. Estes mesmo autores afirmam que a sua popularidade deve-se ao fato deste modelo apresentar uma grande variedade de formas, porém todas com uma propriedade básica: a sua função de taxa de falha é monótona, isto é, ou ela é crescente, decrescente ou constante. Na área de engenharia, a distribuição de Weibull é a mais utilizada para analisar tempo de vida útil de equipamento (ver por exemplo Accioly 1995, Martins e Accioly 2000)

Para uma variável aleatória T com distribuição de Weibull, tem-se a função de densidade de probabilidade dada por:

$$f(t) = \frac{\beta}{\alpha^\beta} t^{\beta-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\beta\right], \quad t > 0 \text{ e } \alpha \text{ e } \beta > 0 \quad (2.10)$$

Se $\beta = 1$, temos a distribuição exponencial.

Para a distribuição Weibull as funções de sobrevivência e de risco para $t > 0$ são, respectivamente:

$$S(t) = \exp\left(-\left(\frac{t}{\alpha}\right)^\beta\right). \quad (2.11)$$

e

$$h(t) = \left(\frac{\beta}{\alpha^\beta}\right) t^{\beta-1} . \quad (2.12)$$

sendo que $h(t)$ é:

- Estritamente crescente para $\beta > 1$;
- Estritamente decrescente para $\beta < 1$;
- Constante para $\beta = 1$ (exponencial).

A média e a variância para esta distribuição são respectivamente:

$$E(T) = \alpha \Gamma\left[1 + \left(\frac{1}{\beta}\right)\right] \text{ e } \text{Var}(T) = \alpha^2 \left\{ \Gamma\left[1 + \left(\frac{2}{\beta}\right)\right] - \Gamma\left[1 + \left(\frac{1}{\beta}\right)\right]^2 \right\}, \quad (2.13)$$

sendo a função gama, $\Gamma(k)$, definida por $\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx$. Os percentis para este modelo são dados por:

$$t_p = \alpha \left[-\log(1-p) \right]^{\frac{1}{\beta}} \quad (2.14)$$

É importante enfatizar aqui, uma relação existente entre a distribuição Weibull e a distribuição valor extremo.

Fazendo $Y = \log(T)$, temos que Y tem uma distribuição Valor Extremo com a seguinte função de densidade.

$$f_y(y) = \frac{1}{\sigma} \exp\left\{ \frac{y-\delta}{\sigma} - \exp\left(\frac{y-\delta}{\sigma}\right) \right\}, \quad y \in \mathfrak{R} \quad (2.15)$$

Sendo $\delta = \log \alpha$ e $\sigma = 1/\beta$. Se $\delta = 0$ e $\sigma = 1$ tem-se a distribuição do valor extremo padrão (Maiores detalhes sobre esta distribuição consultar Lawless (1982)).

A média e a variância da distribuição valor extremo são representadas, respectivamente por:

$$(\delta - \gamma\sigma) \text{ e } \left(\frac{\pi^2 \sigma^2}{6}\right), \text{ onde } \gamma = 0,5772\dots (\text{constante de Euler}) \quad (2.16)$$

As funções de sobrevivência, a taxa de falha e o percentil 100p% da variável Y que tem distribuição valor extremo são dados, respectivamente, por:

$$S(y) = \exp \left\{ -\exp \left\{ \frac{y - \delta}{\sigma} \right\} \right\}. \quad (2.17)$$

$$\lambda(y) = \frac{1}{\sigma} \exp \left\{ \frac{y - \delta}{\sigma} \right\}. \quad (2.18)$$

e

$$t_p = \delta + \sigma \log \left[-\log (1 - p) \right]. \quad (2.19)$$

em que δ é o parâmetro de locação e σ o parâmetro de escala.

2.5 - INFERÊNCIA EM MODELOS PARAMÉTRICOS

Os modelos probabilísticos são caracterizados por quantidades (valores) desconhecidas, chamadas de parâmetros. Em cada estudo envolvendo análise de confiabilidade, os parâmetros devem ser estimados a partir de observações amostrais, para que o modelo fique determinado e, assim, seja possível responder às perguntas de interesse. Entre os vários métodos de estimação conhecidos na literatura estatística, um dos mais conhecidos é o método de máxima verossimilhança.

2.5.1 - O MÉTODO DE MÁXIMA VEROSSIMILHANÇA

Ao determinarmos o modelo probabilístico adequado ao conjunto de dados em análise é desejável fazer inferência baseada nas estimativas dos parâmetros no modelo. Um dos métodos utilizados para obter tais estimativas é o método de máxima verossimilhança. O método funciona da seguinte forma: baseado nos resultados obtidos pela amostra, qual é a distribuição entre todas aquelas definidas pelos possíveis valores de seus parâmetros, com maior possibilidade de ter gerado tal amostra? Em outras palavras, se, por exemplo, a distribuição de falha é a Weibull, para cada combinação diferente de α e β têm-se

diferentes distribuições de Weibull. O estimador de máxima verossimilhança escolhe aquele par de α e β que melhor explique a amostra observada (Colosimo e Giolo, 2006).

Suponha uma amostra de observações t_1, t_2, \dots, t_n de uma certa população de interesse com variável aleatória T . A população é caracterizada pela sua função de densidade de probabilidade $f(t, \theta)$ com $\theta \in \Theta$, sendo Θ é o espaço paramétrico. A função de verossimilhança para um parâmetro genérico θ , correspondendo a uma amostra aleatória é:

$$L(t_1, t_2, \dots, t_n; \theta) = L(\theta) = \prod_{i=1}^n f(t_i; \theta) \quad (2.20)$$

O estimador de máxima verossimilhança de θ é o valor de $\hat{\theta} \in \Theta$ que maximiza a função de verossimilhança $L(\theta)$.

Se todas as observações são não-censuradas, a função de verossimilhança $L(\theta)$ é como dada em (2.20), e mostra que a contribuição de cada observação não-censurada é sua função de densidade $f(t)$.

Contudo, quando existem observações censuradas este método torna-se mais complicado, pois essas informam que o tempo de não sobrevivência é maior que o tempo observado (censurado), sendo portanto sua contribuição para $L(\theta)$ dada pela função de sobrevivência $S(t)$.

Para que se tenha uma notação mais prática, pode-se introduzir uma variável indicadora de censura δ_i , onde $\delta_i = 1$ se t_i é observado diretamente e $\delta_i = 0$ se t_i é censurado à direita, dessa forma, a função de verossimilhança é escrita como (maiores detalhes ver Lawless, 1982)

$$\begin{aligned} L(\theta) &= \prod_{i=1}^r f(t_i; \theta) \prod_{i=r+1}^n S(t_i; \theta) \\ &= \prod_{i=1}^n [f(t_i; \theta)]^{\delta_i} [S(t_i; \theta)]^{1-\delta_i} \end{aligned} \quad (2.21)$$

O processo de maximização torna-se mais simples quando trabalhamos com o logaritmo da função de verossimilhança. Os estimadores de máxima verossimilhança são os valores de θ que maximizam $L(\theta)$ ou equivalente $\log(L(\theta))$. Eles são encontrados resolvendo o sistema de equações

$$U(\theta) = \frac{\partial \log L(\theta)}{\partial \theta} = 0 \quad (2.22)$$

Como citado por Cavalcanti et al. (2002), somente após ter definido qual o modelo que melhor se ajusta aos dados é que deve-se utilizar o método de máxima verossimilhança para estimar os parâmetros μ , σ e os β 's (os coeficientes do modelo). Entretanto, se o modelo selecionado for usado inadequadamente para certo conjunto de dados, toda a análise estatística fica comprometida e conseqüentemente as respostas às perguntas de interesse distorcidas.

Segundo Bolfarine et al. (1991) a forma mais simples e eficiente de selecionar o “melhor” modelo a ser usado para um conjunto de dados é por meio de técnicas gráficas. Este método consiste em fazer gráfico com os possíveis modelos apropriados aos dados e escolher o modelo com comportamento aproximadamente linear.

O método gráfico para o ajuste do melhor modelo paramétrico é apresentado de duas maneiras: A primeira consiste na comparação da função de sobrevivência do modelo possivelmente adequado com a função de sobrevivência obtida através do estimador de Kaplan-Meier. Neste procedimento, o modelo adequado é aquele em que sua curva de sobrevivência se aproxima daquela estimada através do estimador de Kaplan-Meier. A segunda maneira gráfica de analisar o melhor modelo é através da linearização da função de sobrevivência, que consiste em fazer gráficos nos quais o modelo apropriado seja aproximadamente linear. Isto é, o modelo escolhido como adequado para o conjunto de dados será aquele que apresente como resultado final uma reta. Por exemplo, se o modelo Weibull for adequado, o gráfico $\log[-\log(\hat{S}(t))]$ versus $\log(t)$ irá resultar em uma linha reta, passando pela origem (0).

2.6 - AJUSTE DOS DADOS A UMA DISTRIBUIÇÃO EXPONENCIAL E WEIBULL

A seguir uma breve descrição dos estimadores de máxima verossimilhança para os parâmetros dos modelos probabilísticos utilizados nesta dissertação (maiores detalhes ver Collet, 1994).

2.6.1 - DISTRIBUIÇÃO EXPONENCIAL

Suponha uma amostra n , os quais r ($r \leq n$) poços falharam e os $(n - r)$ poços não falharam, isto é, são dados censurados. Considere que os tempos de falhas para os r poços, t_1, t_2, \dots, t_r , tenha uma distribuição exponencial com função de densidades dada pela expressão em (2.7) e função de sobrevivência dada em (2.8). Assim, se substituirmos estas expressões em (2.21), a função de verossimilhança para as n observações dada é

$$L(\alpha) = \prod_{i=1}^n \left[\frac{1}{\alpha} \exp \left\{ - \left(\frac{t_i}{\alpha} \right) \right\} \right]^{\delta_i} \left[\exp \left\{ - \left(\frac{t_i}{\alpha} \right) \right\} \right]^{1-\delta_i} = \prod_{i=1}^n \left[\frac{1}{\alpha} \right]^{\delta_i} \exp \left\{ - \left(\frac{t_i}{\alpha} \right) \right\}$$

onde $\delta_i = 0$ se os tempos são censuras e $\delta_i = 1$ se tempos correspondem a falhas.

O logaritmo da função de verossimilhança é

$$\log L(\alpha) = \log \prod_{i=1}^n \left[\frac{1}{\alpha} \right]^{\delta_i} \exp \left\{ - \left(\frac{t_i}{\alpha} \right) \right\} = - \sum_{i=1}^n \delta_i \log \alpha - \frac{1}{\alpha} \sum_{i=1}^n t_i$$

sabendo que os dados contêm r tempos de falhas, isso indica que $\sum_{i=1}^n \delta_i = r$, então a expressão acima passa a ser

$$\log L(\alpha) = r \log \alpha - \frac{1}{\alpha} \sum_{i=1}^n t_i$$

O objetivo do desenvolvimento das expressões acima é obter uma estimativa de α para a qual o logaritmo da função de verossimilhança seja máximo. Derivando a última expressão em relação α e igualando a zero tem-se

$$\frac{\partial \log L(\alpha)}{\partial \alpha} = -\frac{r}{\alpha} + \frac{\sum_{i=1}^n t_i}{\alpha^2} = 0 \Rightarrow \hat{\alpha} = \frac{\sum_{i=1}^n t_i}{r}. \quad (2.23)$$

Obtendo assim, o estimador de máxima verossimilhança para a distribuição exponencial

2.6.2 - DISTRIBUIÇÃO WEIBULL

Considere novamente uma amostra composta por n poços os quais, r ($r \leq n$) poços falharam. Suponha que os tempos de falhas seguem uma distribuição Weibull com parâmetros de escala α e de forma β . Ao substituir as expressões (2.10) e (2.11) na expressão (2.21) para obter a função de verossimilhança da amostra, a expressão fica dada por

$$L(\beta, \alpha) = \prod_{i=1}^n \left[\frac{\beta}{\alpha^\beta} t_i^{\beta-1} \exp \left[- \left(\frac{t_i}{\alpha} \right)^\beta \right] \right]^{\delta_i} \left[\exp \left(- \left(\frac{t_i}{\alpha} \right)^\beta \right) \right]^{1-\delta_i} = \prod_{i=1}^n \left[\frac{\beta}{\alpha^\beta} t_i^{\beta-1} \right]^{\delta_i} \exp \left\{ - \left(\frac{t_i}{\alpha} \right)^\beta \right\}$$

onde $\delta_i = 0$ se o i -ésimo tempo é censurado e $\delta_i = 1$ caso contrário. Nota-se que a expressão acima é considerada como função de dois parâmetros desconhecidos, α e β . O logaritmo da correspondente função de verossimilhança é dado por

$$\log L(\beta, \alpha) = \sum_{i=1}^n \delta_i \log \left(\frac{\beta}{\alpha^\beta} \right) + (\beta - 1) \sum_{i=1}^n \delta_i \log t_i - \frac{1}{\alpha^\beta} \sum_{i=1}^n t_i^\beta$$

Usando a relação $\sum_{i=1}^n \delta_i = r$, a função torna-se :

$$\begin{aligned} \log L(\beta, \alpha) &= r \log \left(\frac{\beta}{\alpha^\beta} \right) + (\beta - 1) \sum_{i=1}^n \delta_i \log t_i - \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\beta \\ &= r \log(\beta) - r\beta \log(\alpha) + (\beta - 1) \sum_{i=1}^n \delta_i \log t_i - \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\beta \end{aligned}$$

Derivando a expressão acima em relação a β e α e igualando a zero, tem-se

$$\frac{\partial \log L(\beta, \alpha)}{\partial \alpha} = -\frac{r\beta}{\alpha} + \frac{\beta}{\alpha} \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\beta = 0 \quad (2.24)$$

e

$$\frac{\partial \log L(\beta, \alpha)}{\partial \beta} = \frac{r}{\beta} - r \log(\alpha) + \sum_{i=1}^n \delta_i \log(t_i) + \alpha^{-\beta} \log(\alpha) \sum_{i=1}^n (t_i^\beta) - \alpha^{-\beta} \sum_{i=1}^n (t_i^\beta) \log(t_i) = 0 \quad (2.25)$$

As expressões acima serão determinadas aplicando um procedimento numérico iterativo. Geralmente, o procedimento iterativo utilizado para solucionar este tipo de sistema é o algoritmo de Newton- Raphson. Este procedimento consiste em determinar o par de valores $\hat{\beta}$ e $\hat{\alpha}$ que maximizem a função de verossimilhança. Uma vez determinado os valores para $\hat{\beta}$ e $\hat{\alpha}$, a distribuição Weibull estará ajustada aos dados amostrados. Este método iterativo já vem implementado em softwares computacional, portanto não há necessidade de descrever detalhadamente (Maiores detalhes sobre este procedimento consultar Collet (1994) ou apêndice D de Colosimo e Giolo, 2006)

2.7 - MODELO DE REGRESSÃO WEIBULL

Frequentemente, os estudos com dados de tempo de vida apresentam-se relacionados com informações que podem estar influenciando o tempo de sobrevivência. Por exemplo, o tempo até a ocorrência de falha do poço, o evento de interesse nesse estudo, pode depender da profundidade da bomba, do método de elevação, do valor do BSW (*Basic Sediments and Water*) etc. Essas informações são denominadas de covariáveis (também chamadas de variáveis explicativas ou regressoras). As covariáveis podem mudar ou não com o tempo. Quando estão presentes em um estudo é possível verificar a sua influencia na variável resposta.

Uma forma de explicar se uma variável resposta (Y) esta associada com as variáveis explicativas (X_1, X_2, \dots, X_p) é através de um modelo de regressão linear (Hoffmann, 1998).

Quando se tem, como na situação de interesse deste trabalho, tempo até a ocorrência de falha ou censura, e deseja-se utilizar um modelo de regressão para estudar a relação entre elas, o tipo de resposta e o comportamento das variáveis não permitem usar o modelo de regressão linear usual. Além disso, a resposta (tempo de vida) tem em geral uma distribuição assimétrica na direção dos maiores tempos, isso torna inadequado o uso da distribuição normal para o componente estocástico do modelo.

Na literatura existem diversos modelos que descrevem bem o comportamento de dados de confiabilidade, tais como modelo exponencial, Log-Normal, Gama (ver Lawless, 1982 e Colosimo e Giolo, 2006). Em confiabilidade, a maioria das aplicações são baseadas na distribuição Weibull, devido a facilidade do seu uso com dados censurados e pela forma

da sua função risco. Utilizar a distribuição Weibull no contexto da modelagem de tempo de vida significa dizer que o tempo T segue uma distribuição Weibull.

Suponha que T seja uma variável aleatória que mede o tempo até a ocorrência de um evento de interesse e x um vetor composto de $(p+1)$ observações de variáveis explicativas. Denotamos aqui x' como o vetor transposto x . O modelo de regressão Weibull é definido da seguinte forma:

$$Y = \log(T) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p + \sigma v = x' \beta + \sigma v, \quad (2.26)$$

em que $x' = (1, x_1, x_2, \dots, x_p)$ e $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ e T deve ter uma distribuição Weibull para que $\log(T)$ tenha uma distribuição valor extremo com parâmetro de escala σ e locação $\mu(x) = x' \beta$ dependendo das covariáveis. Observe que aqui o erro v tem distribuição valor extremo padrão com densidade $f(v) = \exp\{v - e^v\}$

A função de sobrevivência para Y condicionado a x é expressa por:

$$S(y/x) = \exp\left\{-\exp\left\{\frac{y - x' \beta}{\sigma}\right\}\right\}, \quad (2.27)$$

e para T condicionado a x , por:

$$S(t/x) = \exp\left\{-\left(\frac{t}{\exp\{x' \beta\}}\right)^{1/\sigma}\right\} \quad (2.28)$$

e a função risco expressa por:

$$h(t/x) = \frac{1/\sigma}{\left[\exp(x' \beta)\right]^{1/\sigma}} \cdot (t)^{1/\sigma - 1}$$

Admitindo que o modelo de regressão Weibull é apropriado para o estudo, dá-se início à escolha de um conjunto de variáveis regressoras que melhor expliquem a resposta (ou o tempo de vida). Esta seleção das variáveis explicativas é baseada na comparação de modelos, contendo a presença de covariáveis envolvidas na análise, a fim de selecionar aquele que se ajuste aos dados.

2.8 - ESTRATÉGIA PARA A SELEÇÃO DE COVARIÁVEIS NO MODELO

Freqüentemente em análise de dados de falhas trabalha-se com um conjunto de covariáveis potencialmente importante para descrever o comportamento da resposta. Entretanto, é impraticável ajustar todos os possíveis modelos a fim de selecionar o que melhor explique a resposta. A literatura apresenta algumas técnicas de escolhas do melhor modelo para a variável resposta. Neste estudo, o método de comparação entre modelos de regressão foi baseado no teste da Razão de Verossimilhança (Bolfarine e Sandoval, 2001), e na proposta de seleção sugerida por Collet (1994). A proposta admite que todas as covariáveis apresentam a mesma importância no modelo, portanto aconselha iniciar ajustando um modelo com uma única variável. Os passos utilizados no processo são dados abaixo:

1° passo - Ajustar um modelo para cada uma das covariáveis. Em seguida comparar o valor de cada logaritmo da verossimilhança ($\log L(t)$) obtido com o modelo sem covariável (modelo nulo). Incluir no modelo todas as variáveis que forem significativas ao nível de significância em torno de 10%.

2° passo - Ajustar um modelo com todas as covariáveis significativas no passo 1. Em seguida ajustam-se modelos reduzidos, excluindo uma única covariável de cada vez, examinando quais covariáveis apresentam significância, somente aquelas que apresentarem significância serão importantes para explicar a resposta e permanecem no modelo.

3° passo - As covariáveis incluídas no passo anterior com as covariáveis que não foram significativas no primeiro passo são ajustadas conjuntamente. Neste passo, observa-se se as variáveis excluídas no passo 1 tornaram-se importantes na presença de outras. Existindo interações, estas devem ser incluídas e testadas neste passo.

4° passo - O passo anterior fornece um possível modelo para efeito principal. Para finalizar a modelagem deve-se testar a retirada de cada variável individualmente excluindo do modelo aquelas que não forem significantes. O modelo final fica determinado quando o modelo contiver somente variáveis significantes.

A Tabela 2.1 exemplifica os passos anteriormente descritos, com base em dados hipotéticos. As variáveis são consideradas significativas quando apresentaram p-valor inferior a 0,10.

Tabela 2.1 - Ilustração dos procedimentos citados

Passos	Variáveis no Modelo	Log L(θ)	ξ_{RV} (Estatística do Teste)	Valor-p
Passo 1	Modelo Nulo	- 335,540	-	-
	X₁	- 330,235	10,61	0,0212
	X₂	- 332,715	5,65	0,0933
	X ₃	- 333,756	3,568	0,1816
	X ₄	- 333,599	3,882	0,1635
	X₅	- 329,746	11,588	0,0161
Passo 2	X ₁ + X ₂ + X ₅	- 304,038	-	-
	X ₂ + X ₅	- 305,287	2,498	0,2639
	X₁ + X₂	- 304,165	0,254	0,0667
	X ₁ + X ₅	- 307,398	6,720	0,7226
Passo 3	X ₁ + X ₂	- 304,165	-	-
	X ₁ + X ₂ + X ₃	- 306,529	4,728	0,1619
	X₁ + X₂ + X₅	-320,594	32,856	0,0003
Passo 4	X ₁ + X ₂ + X ₅	-320,594		
	X₁ + X₂	-306,772	26,456	0,0261
	X₁ + X₅	-332,592	24,854	0,0859
	X₂ + X₅	-312,484	16,820	0,0021
Modelo final	X ₁ + X ₂ + X ₅			

3. CONCEITOS FUNDAMENTAIS DE POÇOS DE PETRÓLEO

Este capítulo apresenta, de forma resumida, alguns conceitos sobre engenharia de petróleo, incluindo os conceitos de poços de produção, método de elevação, entre outras. Informações mais detalhadas podem ser obtidas em Thomas (2001).

3.1 - CONCEITOS PRELIMINARES

O petróleo é uma mistura de compostos, cujos principais constituintes são moléculas de hidrocarbonetos. Acredita-se que sua origem esteja ligada a decomposição dos organismos que compõem o plâncton. Estes seres decompostos foram, há milhões de anos, se acumulando no fundo dos mares e dos lagos, sendo pressionados pelos movimentos da crosta terrestre e transformaram-se na substância oleosa que é o petróleo (Apostilas Petrobras).

O petróleo, após ser gerado e ter migrado da rocha geradora, é eventualmente acumulado em uma rocha denominada de *reservatório*. Esta rocha, independentemente da sua origem ou natureza, para se constituir em um reservatório deve apresentar porosidade (espaços vazios no seu interior), permeabilidade (interconexão entre poros). Desse modo, podem se constituir rochas-reservatório os arenitos, calcarenitos e todas as rochas sedimentares essencialmente dotadas de porosidade intergranular que sejam permeáveis. Algumas rochas, como os folhelhos e alguns carbonatos, normalmente porosos, porém impermeáveis, podem vir a se constituir reservatórios quando se apresentam naturalmente fraturados (Thomas 2001).

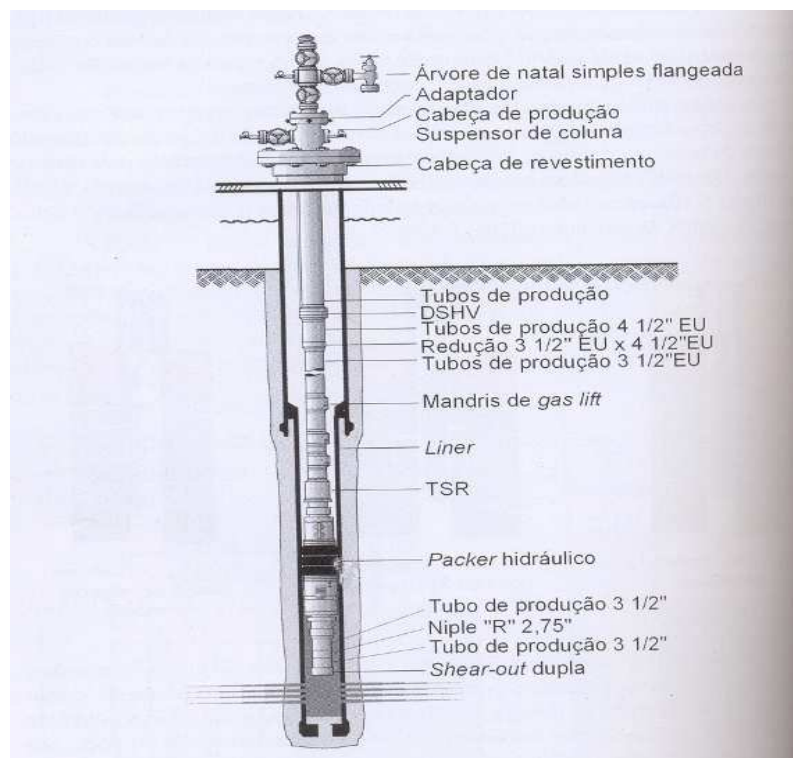
Após a descoberta da rocha com acumulação de petróleo, é necessário realizar estudos geológicos e geofísicos para localizar a jazida (a rocha cujos poros são ocupados pelo petróleo). Esta avaliação determina as zonas da rocha em que deve haver petróleo e se estas são economicamente viável, caso contrário não se explorará a área. Nos casos

economicamente atrativos, realiza-se a perfuração de poços de desenvolvimento na região, formando assim um *campo de petróleo*.

O petróleo pode se apresentar no estado líquido (óleo cru), sólido ou gasoso (gás natural). Os tipos de fluidos mais frequentemente encontrados nos poços terrestres são os fluidos líquidos ou gasosos e a água, que podem variar de estado, dependendo das condições existentes no poço. Neste estudo, os poços alvos são os poços terrestres produtores de óleo da Bacia Potiguar.

Após a perfuração do poço (atividade realizada através de uma *sonda* composta por guincho, torre, mesa rotativa, etc) é necessário fazer a comunicação entre o reservatório e a superfície, isto é, preparar o poço com técnicas apropriadas para que sua produção atinja a superfície. Esta etapa é conhecida como *completação*. Segundo Thomas (2001), esta etapa busca otimizar a vazão e minimizar o risco de futuras intervenções. Quando o poço é totalmente equipado, torna-se um sistema bastante complexo e fica composto pelos seguintes elementos: Equipamentos da superfície: Árvore de natal, Adaptador, Cabeça de produção, Suspensor de coluna e Cabeça de revestimento

Figura 3.1 - Coluna convencional de produção equipada com gás-lift.



Fonte: Thomas, 2001

Segundo Thomas (2001) as colunas de produção tem as seguintes finalidades:

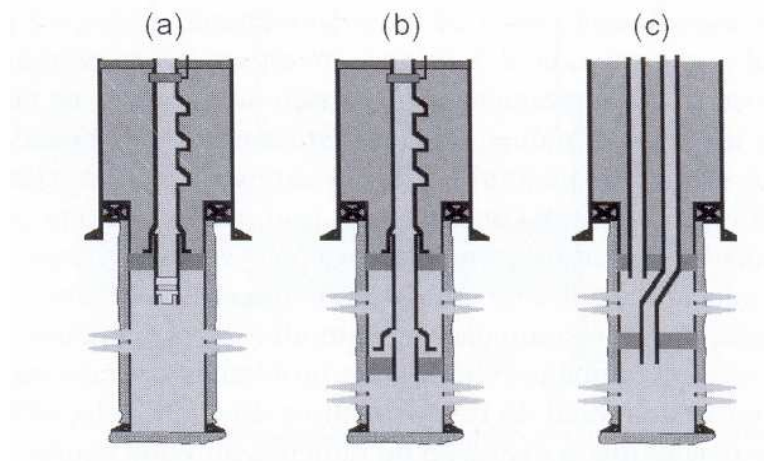
- Conduzir os fluidos produzidos até a superfície, protegendo o revestimento contra fluidos agressivos e pressões elevadas;
- Permitir a instalação de equipamentos para a elevação artificial;
- Possibilitar a circulação de fluidos para o amortecimento do poço, em intervenções futuras.

Ainda segundo este autor a coluna de produção pode ser simples ou múltipla, como mostra a Figura 3.2.

a) *simples*: ocorre quando uma única tubulação metálica é descida no interior do reservatório. Este tipo de completação possibilita produzir de modo controlado e independente apenas uma zona de interesse (Figura 3.2 a). Entretanto mais de uma zona pode ser produzida em conjunto.

b) *Composta*: Permite produzir ao mesmo tempo duas ou mais zonas ou reservatórios diferentes, através de uma ou mais colunas de produção descidas no poço. A Figura 3.2 b mostra um esquema de completação seletiva, com apenas uma coluna de produção, equipada de forma a permitir produção de duas zonas em conjunto ou uma das zonas alternativamente. A Figura 3.2 c mostra um exemplo de completação dupla, que permite a produção de duas zonas de modo controlado e independente, através de duas colunas de produção.

Figura 3.2 - Tipos de completação: (a) simples, (b) seletiva e (c) dupla



Fonte: Thomas, 2001

As colunas de produção múltiplas apresentam maior possibilidade de apresentar problemas operacionais, visto que as instalações são mais complexas, pois há dois sistemas

de elevação independentes. No presente estudo, a coluna de produção será tomada como a unidade de estudo, isto significa dizer que poços equipados com coluna de produção dupla serão analisados como dois poços diferentes.

A grande maioria destes poços não apresenta energia suficiente no fundo do poço para elevar os fluidos até a superfície. Necessitando então, de mecanismo instalado na coluna de produção para fornecer energia para conduzir o fluido até a estação coletora. Alguns poços produzem energia suficiente para produzir, estes são denominadas de poços surgentes. Na engenharia de petróleo existem vários métodos de elevação e a escolha do método depende de algumas características do poço, como por exemplo a profundidade do poço. Os métodos mais utilizados em todo o Brasil são: Bombeio Mecânico (BM), Bombeio por Cavidade Progressiva (BCP), Bombeio Centrifugo Submerso (BCS) e o Gás Lift (GL). Segundo Costa (2004), na quantidade total de poços produtores no mundo, 80% são de bombeio mecânico e ocupa a posição de mais utilizado no mundo. Isto deve-se ao baixo custo de investimento e manutenção, boa eficiência energética e a possibilidade de operar com fluidos de diferentes composições e viscosidades em uma larga faixa de temperatura.

A seguir apresentamos, de forma resumida, os dois métodos de elevação abordados neste estudo, BM e BCP, algumas características e vantagens de cada método (maiores detalhes pode ser visto em Thomas, 2001 ou Miura, 1988).

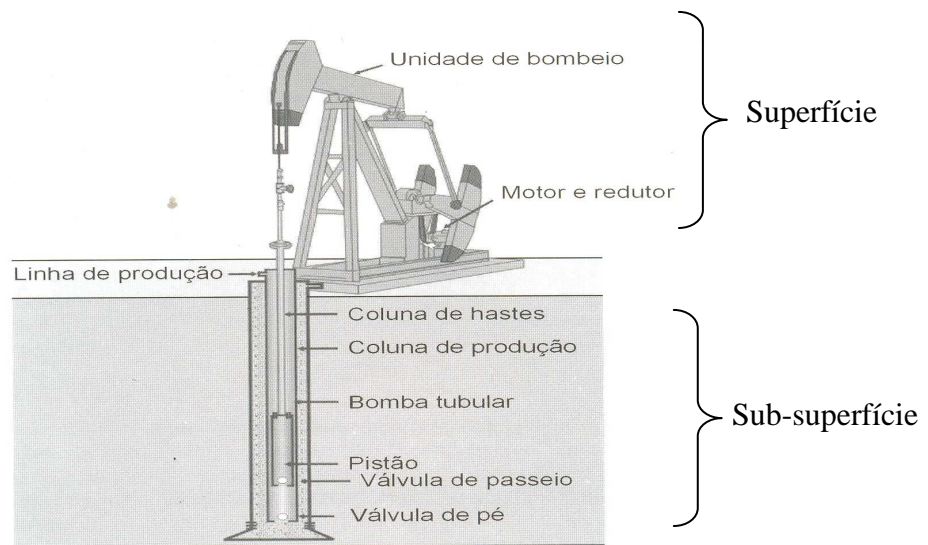
3.2 - BOMBEIO MECÂNICO (BM)

O bombeio mecânico é um método de elevação artificial cujo movimento rotativo de um motor elétrico ou de combustão interna é transformado em movimento alternativo por uma unidade de bombeio localizada próxima à cabeça do poço. Uma coluna de hastes transmite o movimento alternativo para o fundo do poço, acionando uma bomba que eleva os fluidos produzidos pelo reservatório para a superfície (Thomas, 2001).

Segundo Serra (1990) as principais vantagens do BM são: simplicidade de operação, manutenção e projeto de novas instalações; a partir das condições normais pode ser utilizado até o fim da vida produtiva do poço e a capacidade de bombeio pode ser modificada, em função das mudanças de comportamento do poço. Mas, a principal vantagem é o baixo custo de produção ao longo da vida produtiva do poço.

O BM é dividido em dois grupos de componentes. Os componente de superfície e os componentes de sub-superfície, conforme mostrado na Figura 3.3. Os componentes de superfície são: motor, a unidade de bombeio e a haste polida. Os componentes instalados no fundo do poço são: a coluna de hastes e a bomba de fundo. Os componentes são interligados com a finalidade de transmitir energia ao fundo do poço e ajudar na elevação dos fluidos.

Figura 3.3 - Sistema de Bombeio Mecânico



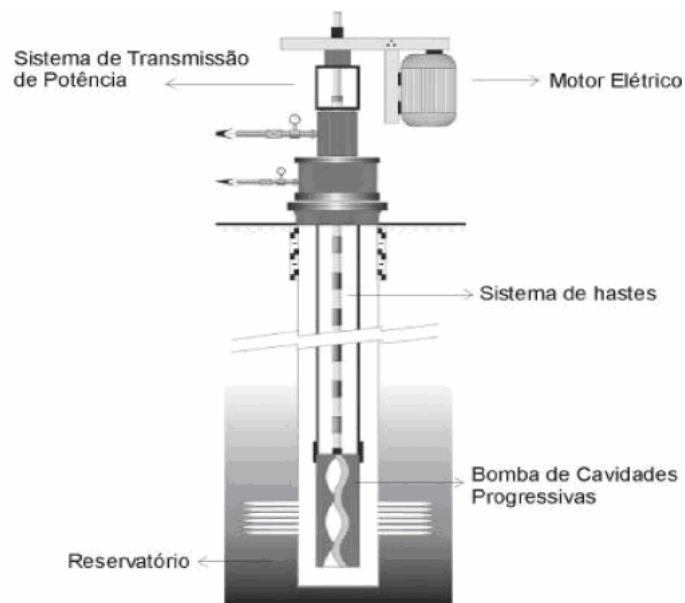
Fonte: Thomas, 2001

3.3 - BOMBEIO POR CAVIDADE PROGRESSIVA (BCP)

O sistema de superfície transmite energia para a bomba de Cavidade Progressiva (BCP) através da coluna de hastes. Esta bomba é composta por rotor e estator. O rotor recebe o movimento de rotação, transmitida pelo motor instalado na cabeça do poço, fazendo-o girar no interior do estator gerando uma série de cavidades idênticas que descolam o fluido da sucção para a descarga. (Nascimento, 2005). A Figura 3.4 apresenta um poço equipado com BCP

De acordo com Thomas (2001), este método é muito utilizado em poço não muito profundo (é usado até 2000 m). Uma das vantagens deste método é que a energia consumida pelo motor é menor do que a consumida pelo BM.

Figura 3.4 - Sistema de Bombeio por Cavidade Progressiva-BCP



3.4 - ACOMPANHAMENTO DA PRODUÇÃO

A vazão de óleo, gás e água, juntamente com a pressão de fundo (ou de nível mecânico) e pressão estática definem a produtividade do poço, ou seja, estabelece a relação entre a vazão e a pressão de fundo. Esta relação é denominada IPR - Inflow Performance Relationship. Define-se por **vazão**, o volume por unidade de tempo, que se escoou através de determinado duto, ou seja é a rapidez com a qual um volume escoou. A unidade adotada pela Petrobras para vazão de poços é m^3 / dia .

Nas medidas de vazão consideradas neste estudo as vazões são expressas nas condições de superfície, também chamadas de “condições padrão”, correspondendo à

pressão de 1 atm¹ e temperatura de 20° C. Por exemplo, temos o metro cúbico standard (m³ /dia).

Produção de óleo

O óleo morto é a parte dos hidrocarbonetos que permanece no estado líquido quando a mistura é levada para a superfície. Quando se diz que um poço está produzindo com uma vazão 100m³ std/dia de óleo, isto significa que da mistura líquida que está saindo diariamente do reservatório através daquele poço, 100 m³ permanecem no estado líquido na superfície.

Produção de gás

O gás nas condições padrão produzido é o resultado da composição de três partes. Uma parte é proveniente dos hidrocarbonetos que, nas condições de temperatura e pressão do reservatório, já se encontram no estado gasoso que tem o nome de gás livre. A segunda parte é o gás que sai de solução do óleo, isto é, os hidrocarbonetos que se encontram dissolvidos no óleo nas condições do reservatório e se vaporizam quando a mistura é elevada para as condições de superfície. A terceira parte é o gás que se encontra dissolvido na água nas condições do reservatório.

Produção de água

Além dos hidrocarbonetos, é bastante comum a produção de água. A quantidade de água produzida vai depender das condições em que ela se apresenta no meio poroso. Apesar da água estar sempre presente nos reservatórios, nem sempre a sua quantidade, expressa pela sua saturação, é suficiente para que ela se desloque. Existe uma saturação mínima de água a partir da qual ele se torna móvel. Essa saturação depende da rocha e dos fluidos nela contidos. Se a saturação de água for igual a esse valor mínimo, não haverá fluxo, e conseqüentemente não haverá produção de água dessa rocha (Thomas, 2001).

Existem algumas relações dentro da engenharia de petróleo que são utilizadas como indicadores, tanto de características como de estágios da vida produtiva dos reservatórios.

¹ Pressão atmosférica é a pressão exercida sobre a superfície da Terra pelo peso da atmosfera

Os mais utilizados são a razão gás-óleo, RGO, a razão água-óleo, RAO, e o BSW (do inglês *basic sediments and water*).

RGO: Frequentemente um campo produz ao mesmo tempo petróleo e gás natural, havendo, em alguns, a predominância do petróleo e, noutros, a do gás natural. Para medir qual dos dois predomina, é utilizado um coeficiente chamado razão gás-óleo, conhecido como RGO. A razão gás-óleo é a relação entre a vazão de gás e a vazão de óleo; ambas medidas nas condições de superfície.

RAO: a razão água-óleo é a relação entre a vazão de água e a vazão de óleo, ambas medidas nas condições de superfície. Uma razão água-óleo igual a zero significa que, na época da medição, a saturação de água na zona de onde está saindo a mistura de hidrocarbonetos é igual ou menor que o valor irreduzível (Thomas, 2001).

BSW: é o quociente entre a vazão de água mais os sedimentos que estão sendo produzidos e a vazão total de líquidos e sedimentos. ($BSW = \text{vazão de (água + sedimentos)} / \text{vazão de (óleo + água + sedimentos)}$)

3.5 - OPERAÇÕES DE MANUTENÇÃO EM POÇOS

Mesmo após a sua completação, os poços de petróleo irão necessitar de algumas intervenções para manter a produção ou até mesmo para melhorar seu desempenho. Estas intervenções são conhecidas na engenharia do petróleo como *Workover*.

Segundo Garcia (1997), a manutenção da produção é o conjunto de operações realizadas no poço visando corrigir problemas de modo a permitir que a produção de fluidos retorne ao nível normal ou operacional. Um dos motivos para realizar uma intervenção, pode ser, por exemplo baixa produtividade ou produção excessiva de gás. Algumas destas intervenções necessitam do deslocamento de uma sonda, tornando a manutenção muito onerosa. A Petrobras realiza algumas operações para manutenção da produção dos poços, e de acordo com este autor os tipos de operação são:

- **Avaliação:** É um conjunto de operações que tem como objetivo avaliar o potencial do reservatório do petróleo, sua capacidade produtiva e a valorização das suas reservas.

- Restauração: Conjunto de atividades realizadas com o objetivo de devolver as condições normais de entrada de fluido do reservatório para o poço, eliminar e/ou corrigir falhas mecânicas no revestimento ou na cimentação, reduzir a produção excessiva de gás ou de água.
- Limpeza: Conjunto de atividades realizadas no interior do revestimento com o objetivo de retornar ao potencial inicial do poço. Quaisquer falhas em equipamentos ou até mesmo a necessidade de substituição destes são motivos para realizar a limpeza.
- Estimulação: O objetivo da estimulação é aumentar o índice de produtividade de um reservatório.
- Mudança do Método de Elevação (M.M.E): Operação que tem como objetivo substituir um determinado método de elevação por outro. Esta substituição ocorre quando durante a produção, o poço apresentar deficiência na vazão devido a escolha do método de elevação inadequada ou por apresentar algum defeito.
- Abandono: O abandono do poço é realizado quando o poço atingiu o final da vida produtiva ou quando após a perfuração constata-se que o mesmo não é viável comercialmente. O abandono pode ser *temporário* – realizado quando ainda há possibilidade de retorno ao poço - ou *definitivo* que é realizado quando não se prevê o retorno ao poço, isto é, o poço atingiu todo o seu potencial.

Para uma visão geral da engenharia de poço consultar Thomas (2001).

3.6 - DEFINIÇÃO DO ESTADO DOS POÇOS

A Petrobras, através de códigos define o estado dos poços durante sua vida produtiva. O poço pode estar em seis tipos de estados e alguns destes são usados quando o poço não está funcionando. Estes estados são representados por códigos de 1 a 6. Cada código define a situação em que se encontra o poço.

Antes de detalhar cada estado, é importante descrever algumas definições usadas pela Petrobras referentes à produção do poço. Define-se como *produção/injeção-base* a quantidade de fluido produzido pelo poço, medido em m³. O *potencial-base* refere-se ao valor da produção estimada para o poço dentro das condições normais de funcionamento.

Quando ocorre algum problema, seja técnico ou operacional, a produção-base difere do potencial-base, esta diferença é conhecida como **perda**.

Os seis tipos de estados utilizados são:

1° Estado – Código 1 - Poço em operação normal. Refere-se ao poço-coluna que está com sua produção/injeção-base igual ao potencial-base, conseqüentemente com perda-base igual a zero;

2° Estado – Código 2 - Poço operando com deficiência. Refere-se ao poço-coluna que está operando com alguma anormalidade, o que faz produzir (ou injetar) com vazão inferior ao potencial-base. Necessitando de algumas providencias (intervenções ou manutenções) para retornar ao estado normal ou reavaliar o seu potencial-base.

3° Estado – Código 3 - Poço fechado com perda devido a problemas operacionais. Refere-se ao poço-coluna fechado contabilizando perda de produção devido a um problema operacional, isto é, um problema no equipamento vital da coluna de produção que causou o fechamento no poço. Sua produção-base é zero e a perda é igual ao potencial-base. Aguardando providencias para entrar em produção. Neste estudo, a *falha* do poço-coluna será caracterizada quando o poço mudar para este estado e a causa da ocorrência da anormalidade estiver relacionada à falha em um equipamento de subsuperfície do poço.

4° Estado – Código 4 - Poço produtor fechado por estratégia. Refere-se ao poço-coluna fechado temporariamente por decisão estratégica. Não contabiliza perda. Poderá voltar ao estado normal através de custeio de operação ou projetos de investimentos.

5° Estado – Código 5 - Poço-coluna fechado sem perda. Aguardando operação de abandono. A colocação de um poço neste estado requer uma análise técnica e econômica multidisciplinar e deve ser formalizada através de documento a ser definido pela administração da UN-RNCE.

6° Estado – Código 6 - Poço desativado. Refere-se ao poço-coluna desativado. Para essa situação, a coluna do poço foi completamente removida, retirados de operação.

3.7 - CLASSIFICAÇÃO DOS POÇOS

A perfuração de poço tem como finalidade produzir petróleo e/ou gás natural. Cada poço recebe um código de acordo com um regulamento aprovado pela ANP - Agência Nacional de Petróleo Petrobras, de N° 75, de 3 de maio de 2000. Este regulamento trata dos procedimentos a serem adotados pela Petrobras para a codificação de poços perfurados com vistas à exploração ou produção de petróleo e/ou gás. A codificação consiste em dotar o poço de um nome e um cadastro diferente, de forma a não haver duplicidade.

De acordo com este regulamento, o nome do poço é um conjunto de símbolos alfanuméricos que identifica o poço em relatórios, mapas e demais documentos. Já o cadastro é um conjunto de algarismos agrupados de forma que permita a identificação do poço. O nome deve ser composto por cinco partes referentes a: **categoria** - define o poço segundo sua finalidade, **referencial nominal** - estabelece o conjunto de 2 a 4 letras que compõem o nome do poço; **numeração** - numera os poços sequencialmente em ordem de perfuração, **tipo** - define o poço quanto a sua geometria e por **referencia geográfica** - conjunto de letras que identifica a Unidade de Federação onde o poço se localiza. A junção das cinco categorias tem como objetivo identificar, localizar a região, conhecer a geometria e a finalidade do poço.

Os poços podem ser do tipo exploratórios ou explotatório. Os exploratórios têm como finalidade explorar o reservatório, já os explotatórios têm o objetivo de melhorar a produção de um campo de produção.

A seguir uma breve descrição das classificações definidas para os poços seguindo o regulamento da ANP (BRASIL, 2000). Os poços são classificados de acordo com sua finalidade. São eles:

- Poço Exploratório Pioneiro – identificado com o código 1- refere-se ao poço que tem como objetivo testar a ocorrência de petróleo e/ou gás natural em um reservatório em um ou mais objetivos de um prospecto geológico;
- Poço Exploratório Estratigráfico, identificado com o código 2, é aquele perfurado com a finalidade de conhecer-se a coluna stratigráfica (coluna que mostra os limites da rocha onde a perfuração deve atravessar até chegar ao objetivo final) de uma bacia e obter outras informações geológicas de subsuperfície.

- Poço Exploratório de Extensão, identificado com o código 3, é aquele que visa delimitar a acumulação de petróleo e/ou gás natural em um reservatório;
- Poço Exploratório Pioneiro Adjacente, identificado com o código 4, é aquele que visa testar a ocorrência de petróleo e/ou gás natural em uma área adjacente a uma descoberta;
- Poço Exploratório para Jazida Mais Rasa, identificado com o código 5, é aquele que visa testar a ocorrência de jazidas mais rasas do que as já descobertas numa determinada área;
- Poço Exploratório para Jazida Mais Profunda, identificado com o código 6, é aquele que visa testar a ocorrência de jazidas mais profundas do que as já descobertas numa determinada área;
- Poço Exploratório de Produção, identificado com o código 7, é aquele que visa drenar uma ou mais jazidas de um campo;
- Poço Exploratório de Injeção, identificado com o código 8, é aquele destinado à injeção de fluidos visando melhorar a recuperação de petróleo, de gás natural ou a manter a energia do reservatório;
- Poço Especial, identificado com o código 9, é aquele que visa permitir uma operação específica que não se enquadra nas situações anteriormente definidas;

Neste trabalho, a classificação do poço será inclusa na análise para verificar como esta covariável está relacionada com tempo de falha dos poços.

3.8 - SISTEMAS DE INFORMAÇÕES DA PETROBRAS

A Petrobrás dispõe de sistemas de coleta de dados que armazenam diversas informações sobre os poços, desde a sua data de criação até o final de vida produtiva, assim como testes de medições. Os dados utilizados neste estudo encontram-se disponíveis em 2 sistemas: o SIP e SEP descritos a seguir.

SIP – Sistema de Informação de Produção

Estruturado em Oracle, em ambiente corporativo, de âmbito nacional, o SIP é o sistema oficial de informação e controle da produção e injeção de fluidos da Petrobras. O principal objetivo do sistema é manter o registro histórico dos dados do acompanhamento

da produção e injeção e manter informação sobre volumes realizados e não realizados dos poços-colunas da Petrobras. Contém informações como resultadas dos testes de produção dos poços, estado de cada poço, data de início de operação de cada poço, produção bruta e líquida de óleo, vazão de injeção de água, potencial de produção de cada poço, etc.

SEP – Sistema de Engenharia de Poço

Criado pela Petrobrás, estruturado em Oracle, este sistema englobava boletins diários de perfuração, avaliação, completação e relatórios de anormalidades/falhas internas e externas ao poço. É alimentado diretamente pelos executores das operações, ou seja, engenheiros e técnicos especializados em suas respectivas áreas de atuação. As correções, as eventuais inconsistências quanto ao banco de dados, migração de dados e a confiabilidade são verificadas por uma equipe de suporte e apoio ao usuário.

4. PLANO AMOSTRAL PARA SELEÇÃO DOS POÇOS

Neste trabalho consideramos como unidade de medida o sistema poço-coluna, que representa o conjunto de todos os seus componentes e estamos interessados na caracterização da primeira falha em equipamentos de sub-superfície. Consideramos que em um poço-coluna funciona um conjunto de itens em série e que a falha de um único item implica na falha do sistema (poço-coluna). Definimos, então, falha como a interrupção total do funcionamento do poço-coluna devido à falha de um ou mais componentes.

Quando o sistema não apresentou falha até o fim do estudo ou quando foi retirado de operação, a observação foi considerada um tempo censurado (Colosimo e Giolo, 2006).

4.1 - COLETA DE DADOS

Foram coletados dados sobre falhas de poços de petróleo no período de janeiro de 2000 a dezembro de 2006 com o objetivo de desenvolver um esquema de previsão de falhas em equipamentos de sub-superfície na Unidade de Negócios do Rio Grande do Norte e Ceará (UN-RNCE). Os dados levantados incluem informações relevantes apenas dos poços terrestres e produtores de óleo cujo método de elevação seja BM ou BCP.

Os dados foram extraídos das planilhas dos sistemas de registro de informações utilizados pela Petrobras, o SIP e o SEP. Foram analisadas planilhas com variáveis consideradas importantes para análise do tempo de falhas dos poços. Os dados foram cruzados entre si com o objetivo de organizar em uma única tabela informações necessárias para facilitar as consultas e aplicação do modelo selecionado para análise dos dados.

Todas as variáveis foram estudadas, analisadas e tabuladas no Excel 2003 e no software estatístico R (R. Development Core Team, 2007), versão 2.6, com a finalidade de eliminar qualquer erro no conjunto de dados e reproduzir um banco de dados confiável.

Esta fase foi minuciosamente detalhada, pois o sucesso da técnica de análise escolhida está diretamente relacionado com quão confiáveis são os dados.

4.2 - POPULAÇÃO CONSIDERADA NO ESTUDO

População é um conjunto de elementos (na totalidade) que são possíveis de se observar e que tem em comum uma determinada característica, sobre a qual se deseja estabelecer conclusões ou exercer ações.

Em nosso estudo, a população alvo foi composta por 4.058 poços terrestres da Bacia Potiguar sujeitos aos métodos de elevação artificial por bombeio mecânico e bombeio por cavidades progressivas onde foram amostrados alguns poços para coletar informações sobre falhas ocorridas no período determinado e medir características dos poços nesse período.

4.3 - AMOSTRA

Segundo Barbetta (2001), nas pesquisas em que se quer conhecer algumas características de uma população, é muito comum se observar apenas uma amostra de seus elementos e, a partir dos resultados dessa amostra, obter valores aproximados ou estimativas para as características populacionais de interesse.

Para o planejamento do plano amostral, foi realizada uma avaliação criteriosa das variáveis disponíveis nos sistemas de informações da Petrobras. Após esta análise formou-se uma base de dados apenas com as variáveis relevantes para elaboração do plano amostral. A identificação destas variáveis nesta nova base foi a seguinte:

- codpoco: código identificador do poço;
- refpoco: código identificador do campo de produção;
- nrpoco: número do poço no campo de produção;
- ufpoco: unidade da federação (CE ou RN);
- coluna: identificador da coluna de produção no poço (U - única, I - inferior, S- superior);
- prodpot: produção potencial do poço na data da coleta;

- estado: estado de produção do poço (escala de 1 a 6, com 1 normal e 6 desativado);
- prodest: produção estimada do poço;
- intervalo: intervalo de profundidade do poço;
- zona: zona de produção;
- metelev: método de elevação de petróleo do poço;
- campo: código identificador da localização do poço;
- ambiente: tipo de produção do poço (T - terrestre, M - marítima);
- unidadm: unidade administrativa indicando o ativo de produção (ATP-ARG – Alto do Rodrigues e ATP-MO - Mossoró);
- município: município de localização do poço;
- profbomb: profundidade medida da bomba (metros);
- prodagua: produção acumulada de água (m³);
- prodoleo: produção acumulada de óleo (m³);
- prodgas: produção acumulada de gás (m³);

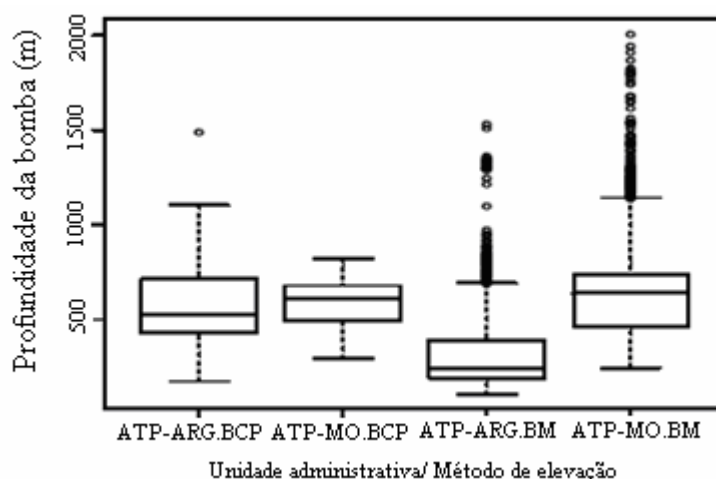
Foram considerados que as unidades amostrais (ou de amostragem) neste estudo seriam as colunas de produção instaladas em um mesmo poço, uma vez que os equipamentos e o funcionamento destas colunas são desconectados e, naturalmente, as falhas ocorrem de modo independente.

No sentido de se garantir que a amostra a ser retirada fornecesse informações para dois ativos de produção (Alto do Rodrigues e Mossoró) e os métodos de elevação BM e BCP, optou-se inicialmente por realizar amostragem estratificada retirando-se amostras separadas de cada estrato dos grupos definidos por estes dois critérios.

Como o objetivo da dissertação recai sobre a medição de tempo da primeira falha nos poços e de possíveis variáveis relacionadas, avaliou-se também a inclusão nos critérios de estratificação de variáveis disponíveis na base de dados com possíveis associações com as variáveis de interesse. Após várias simulações com a profundidade da bomba, produção acumulada de óleo, razão gás líquido e produção potencial, decidiu-se pela inclusão apenas da profundidade da bomba, uma vez que as outras tinham características dinâmicas, ou seja, seus valores poderiam mudar com o tempo. Assim, as variáveis, profundidade da bomba, método de elevação e unidade administrativa foram escolhidas para definir os estratos.

Para verificar a variabilidade da variável profundidade por unidade administrativa e por método de elevação foi construído um box-plot para analisar o comportamento desta variável nos quatro estratos. A Figura 4.1 exibe os diagramas de caixa para as medidas de profundidade por unidade administrativa e por método.

Figura 4.1 - Distribuição das medidas de profundidade da bomba segundo a unidade administrativa e o método de elevação.



Pela Figura 4.1 nota-se que a variabilidade das profundidades variam entre os quatro grupos. Também, verifica-se que os poços mais profundos são em geral submetidos a bombeio mecânico. Portanto, para formar estratos mais homogêneos, subdividiu-se cada grupo de acordo com faixas de profundidade, sendo duas em cada grupo com método de elevação BCP e três faixas em cada grupo com bombeio mecânico, formando assim, um total de dez estratos. A definição final dos estratos é dada na Tabela 4.1.

Tabela 4.1 - Definição e tamanho dos estratos para a população de 4.358 colunas de produção.

Numero de Estratos (h)	Unidade administrativa	Método de elevação	Profundidade (m)	Nº de poços-colunas (N _h)
1	ATP-ARG	BCP	< = 590	70
2	ATP-ARG	BCP	> 590	49
3	ATP-MO	BCP	< = 572	55
4	ATP-MO	BCP	> 572	63
5	ATP-ARG	BM	< = 375	1527
6	ATP-ARG	BM	(375,647]	405
7	ATP-ARG	BM	> 647	161
8	ATP-MO	BM	< = 619	938
9	ATP-MO	BM	(619,995]	874
10	ATO-MO	BM	> 995	216
Total				4.358

4.3.1 - TAMANHO E ALOCAÇÃO DA AMOSTRA

Para a definição do número de colunas “ n ” a serem selecionadas na população de 4.358 colunas, bem como os respectivos números de coluna na amostra de cada estrato (n_h), levou-se em conta que valores grandes de n poderiam dificultar o armazenamento das observações a serem coletadas em cada coluna devido ao elevado número de variáveis relacionadas com os tempos de falha e, possivelmente, as muitas datas em que ocorreram as falhas. Considerou-se um tamanho de amostra $n = 450$ colunas. A alocação final juntamente com os pesos amostrais de cada estrato é exibida na Tabela 4.2.

As Tabelas 4.3 a 4.5 fornecem o número de unidades amostradas em cada unidade administrativa e método de elevação. (Detalhes técnicos sobre este método de estimação ver Anexo I)

Tabela 4.2 - Alocação da amostra e pesos amostrais

Número do Estrato (h)	Unidade administrativa	Método de elevação	Profundidade (m)	N_h	Tamanho da amostra (n_h)	Peso amostral
1	ATP -ARG	BCP	≤ 590	70	16	4,38
2	ATP -ARG	BCP	> 590	49	14	3,50
3	ATP -MO	BCP	≤ 572	55	15	3,67
4	ATP -MO	BCP	> 572	63	16	3,94
5	ATP -ARG	BM	≤ 375	1.527	140	10,91
6	ATP -ARG	BM	(375, 647]	405	37	10,95
7	ATP -ARG	BM	> 647	161	20	8,05
8	ATP -MO	BM	≤ 619	938	86	10,91
9	ATP -MO	BM	(619,995]	874	80	10,93
10	ATP -MO	BM	> 995	216	26	8,31
Total				4.358	450	

Tabela 4.3 - Número de colunas amostradas, por unidade administrativa.

Unidade administrativa	Tamanho do estrato	Tamanho da amostra
ATP -ARG	2.212	227
ATP -MO	2.146	223
Total	4.358	450

Tabela 4.4 - Número de colunas amostradas, por método de elevação

Método de elevação	Tamanho do estrato	Tamanho da amostra
BCP	237	61
BM	4.121	389
Total	4.358	450

Tabela 4.5 - Número de colunas amostradas, por unidade administrativa e método de elevação

Unidade administrativa	Método de elevação	Tamanho do estrato	Tamanho da amostra
ATP -ARG	BCP	119	30
ATP -ARG	BM	2.093	197
ATP -MO	BCP	118	31
ATP -MO	BM	2.028	192
Total		4.358	450

4.4 - SELEÇÃO DA AMOSTRA

O procedimento de amostragem selecionou amostras de cada estrato aleatoriamente, atribuindo-se iguais probabilidades de seleção a cada uma colunas no estrato. O número de colunas sorteadas em cada estrato foi dado pelas quantidades n_h , na Tabela 4.2. A composição das dez amostras produzidas gera uma amostra aleatória estratificada (AAE) de tamanho 450 da população de 4.358 colunas. No sentido de estimar a profundidade média da bomba na população com o estimador de Horvitz-Thompson (1952), a amostra retirada produz uma margem sobre o erro amostral de aproximadamente 8 metros, com grau de confiança de 95%.

5. RESULTADOS

Neste capítulo serão apresentados os resultados obtidos na análise de dados de tempos de falhas dos poços. Inicialmente serão apresentadas as variáveis consideradas importantes para o estudo e uma análise descritiva das mesmas. Esta análise consiste na organização dos dados em tabelas, gráficos e algumas medidas, com intuito de facilitar a compreensão das informações referentes aos poços. Em seguida serão apresentados alguns resultados não paramétricos para estimar as curvas de sobrevivência, a seleção do melhor modelo paramétrico de regressão, as estimativas dos parâmetros e a exibição das covariáveis consideradas significativas para explicar o tempo até a primeira falha dos poços-coluna.

O evento de interesse neste estudo é a *falha* do poço. A variável resposta é o tempo de funcionamento do poço dentro da sua normalidade até apresentar a primeira falha, definida como alguma anormalidade relacionada a equipamentos de subsuperfície que tenha ocasionado uma parada total no funcionamento do poço. A *censura* foi caracterizada pela não ocorrência de falha no poço até o fim do estudo ou pelo fato do poço ter sido desativado ou retirado de operação antes de ocorrer o evento.

Os dados utilizados foram analisados com a ajuda do software computacional R, versão 2.6 e o Excel, versão 2003. Nos anexos encontram-se a descrição das tabelas, as codificações das variáveis e os comandos utilizados para obter os resultados aqui descritos.

5.1 - ANÁLISE EXPLORATÓRIA DOS DADOS

Neste estudo, foram considerados dados provenientes de dois sistemas de coleta de informações utilizados pela Petrobras. Cada sistema apresenta estrutura e informações diferentes. Nenhum deles contém informações completas para a análise de tempo de falhas dos poços. Portanto, fez-se necessário analisar as duas bases e extrair informações

relevantes para o estudo. Esta etapa demandou várias reuniões com técnicos e engenheiros da Petrobras para consolidar o entendimento das variáveis existentes e construir um banco de dados mais enxuto. Após a seleção de algumas variáveis, a fase seguinte foi organizar uma única base de dados contendo todas as informações possivelmente associadas com o tempo de falha.

Como o objetivo do trabalho consiste em estudar as falhas de equipamentos de subsuperfície dos poços de petróleo que utilizam o método de elevação Bombeio mecânico (BM), Bombeio por cavidade progressiva (BCP) e produtores de óleo, foi necessário excluir todas as informações dos poços que em algum momento produziram com os métodos de elevação diferentes de BM e BCP e todas as informações sobre produção de água e gás.

Como descrito na Capítulo 4, a unidade amostral considerada foi a coluna de produção do poço e não o poço em si, visto que o poço pode produzir através de duas colunas independentes. Foi determinado que a amostra mínima, para a representação dos poços da Bacia Potiguar, seria de 450 poços-coluna. Sabe-se que com o decorrer do tempo, alguns poços exigem mudança ou instalação de outra coluna de produção. No período analisado, alguns poços passaram a produzir com um tipo de coluna diferente daquela do início do estudo ou necessitaram do auxílio de mais uma coluna para atender toda a produção do poço. Portanto, considerou -se que uma mudança na coluna de produção representava que uma nova unidade amostral entrou em produção. Conseqüentemente, a amostra passou para 603 pocos-coluna, no período de 2000 a 2006.

As variáveis e alguns códigos identificadores dos poços inclusos na base de dados foram:

- cd_poço : código do poço (código de registro do poço);
- coluna: identificador da coluna de produção no poço (U - única, I - inferior, S- superior);
- dt_inicio: data que o poço iniciou suas operações em condições normais;
- produção base: produção base do poço, medida em m³/dia;
- dt_fim: data que o poço finaliza sua produção em condições normais. A interrupção da produção pode ser devido a problemas técnicos ou por parada programada (visando melhorar a produção).

- dt_ini_oper: data de início de operação prevista ou efetiva para início de operação do poço
- cd_estado: situação da produção do poço (escala de 1 a 6, com 1 normal e 6 desativado);
- cd_fluido: o código identificador da produção do fluido de cada poço (a produção pode ser óleo, gás e água, identificados, respectivamente, pelos códigos: 187, 181 e 173). Neste estudo, serão analisados apenas os produtores de óleo;
- vl_BSW: valor do BSW (Basic Sediments and Water). É a quantidade de água produzida conjuntamente com o petróleo;
- vl_RGO: é a Razão Gás-Óleo (medida em m^3/m^3);
- cd_grupo_prov: é o código da providências de uma ocorrência. É a providência utilizada na correção de alguma anormalidade do poço.
- método de elevação: é o método artificial utilizado para elevar o fluido, contido no fundo do poço, até a superfície do poço. Neste estudo, os dois métodos considerados foram BM – Bombeio mecânico e BCP – Bombeio por Cavidade Progressiva;
- unidade operativa: Unidade operativa do poço. Cada unidade administra os poços de acordo com a sua localização geográfica. São quatro unidades: OP_ARG – Unidade Operacional Alto do Rodrigues, OP_CAM – Unidade Operacional Canto do Amaro, OP_ET – Unidade Operacional Campo de Estreito e OP ATP_RFQ – Unidade Operacional Fazenda Riacho da Forquilha;
- Índice de Produtividade (IP): Valor do índice de produtividade. Este índice é calculado da seguinte forma: $IP = p / (P_e - P_w)$ – onde p é vazão, p_e a pressão estática (ou média) do reservatório e P_w e a pressão de fluxo no fundo do poço;
- fluido previsto: quantidade de fluido previsto para o poço no início da operação;
- localização da UF: Estado da Bacia o qual o poço esta localizado, os dois estados são: CE e RN;
- classificação do poço: Tipo de poço. A classificação do poço é dada de acordo com sua finalidade. É dito exploratório – quando sua finalidade é explorar o reservatório, ou pode explotatório – quando sua função é melhorar a produção de um campo, isto é, a finalidade da perfuração do poço é ajudar na produção dos poços existentes. São eles: Poço exploratório pioneiro, Poço Exploratório Estratigráfico, Poço Exploratório de Extensão, Poço Exploratório Pioneiro Adjacente, Poço Exploratório para Jazida Mais Rasa, Poço Exploratório para

Jazida Mais Profunda, Poço Explotatório de Produção, Poço Explotatório de Injeção e Poço especial.

- *cd_motivo*: Códigos dos motivos utilizados para registrar as ocorrências de um poço-coluna. Isto é, representa a causa da parada para realizar uma ocorrência no poço. Entre outras ocorrências, nesta variável pode-se observar as ocorrências relacionadas equipamentos de subsuperfície. A diferença entre a data de funcionamento normal do poço e a data desta primeira ocorrência corresponde o *tempo até a primeira falha* do poço que é variável reposta de interesse deste trabalho.
- *profundidade do poço*: profundidade do poço medida em metro (m);
- *profundidade da bomba*: profundidade onde se encontra instalado a bomba de produção do poço (m);

Além destas variáveis contidas na base de dados, foram calculadas posteriormente a idade do poço na data da falha e o tempo que o poço levou desde início do estudo até a primeira a falha. Foi construído também nesta planilha um indicador de censura, onde o poço-coluna recebia indicador de censura 1 se o poço-coluna apresentasse falha até a data final do estudo, recebia zero (0) caso contrário.

Os dados organizados e condensados resultaram em uma planilha contendo 603 linhas e 27 colunas com informações de todas as variáveis relevantes e suas respectivas medidas calculadas com os dados do período de 2000 a 2006.

Antes da aplicação do modelo estatístico nos dados realizou-se uma análise exploratória das variáveis e verificou-se que a variável índice de produtividade e Quantidade de fluido previsto não deveriam ser inclusas no modelo por apresentar grande quantidade de dados faltantes. Verificou-se que 83,16% dos poços apresentavam a variável Índice de Produtividade vazias e 50% na variável Quantidade de fluido.

A seguir são apresentadas estatísticas descritivas das variáveis quantitativas consideradas no banco de dados para verificar a possível associação no tempo até a falha do poço.

Tabelas 5.1 - Estatística descritiva das variáveis numéricas relevantes para o estudo

Variáveis	Min	1° Quartil (25%)	Mediana	Média	3° Quartil (25%)	Máx
Produção base	0.10 m ³ /dia	1.10m ³ /dia	2.20m ³ /dia	4.02m ³ /dia	4.80m ³ /dia	65.00 m ³ /dia
Valor do BSW	0.10%	44.60 %	79.40%	67.10%	92.58%	99.90%
Valor do RGO	1.00m ³	1.00m ³	4.00m ³	6.83m ³	9.00m ³	816.90 m ³
Profundidade do poço	235.1 m	397.8 m	675.0m	672.9 m	869.5 m	2160.0 m
Profundidade da bomba	139.0m	307.0 m	490.0 m	522.6 m	692.5 m	1678.0 m
Idade do poço (anos)	0.093	2.39	6.71	7.79	12.85	20.97

Como observado na Tabela 5.1, os poços amostrados da bacia potiguar apresentaram uma produção média de óleo de 4.02 m³/dia. Do total de poços que compõem a amostra, 50% deles produzem diariamente menos de 2.20 m³/dia de óleo.

O BSW médio foi de 67,10% por dia, isto é, em média, 67,10% da produção diária dos poços correspondem à água. Nota-se também, que metade dos poços apresentou uma produção de BSW abaixo de 79,40%. Estas medidas indicam que os poços da bacia potiguar apresentam produção de água muito elevada.

A razão gás óleo média (RGO) dos poços amostrados foi de 6.83m³, mostrando que a produção de gás é mais elevada em relação a produção de óleo.

A profundidade média dos poços foi de 677,4 m e a profundidade média da instalação da bomba era de 525.6m. Os dados mostram também que 50% dos poços têm a bomba instalada em profundidade inferior a 490,0m.

Com relação a idade dos poços, nota-se que a média para os poços em análise era de aproximadamente 8 anos e a idade mediana 7 anos. De acordo com os dados analisados a idade máxima observada foi de aproximadamente 21 anos.

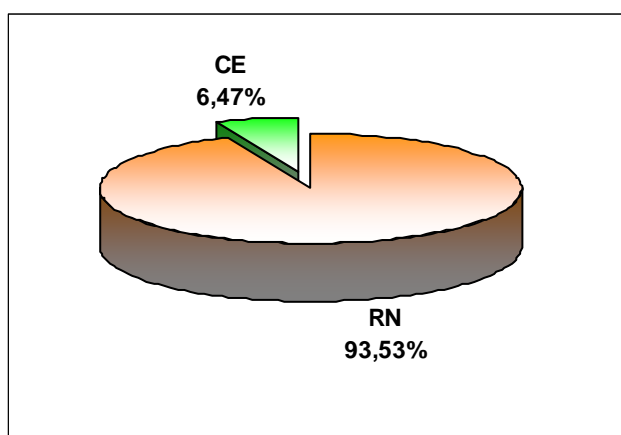
Conforme a Tabela 5.2 e a Figura 5.1, a maioria dos poços selecionados, 564 (93.53%) pertencem ao estado do RN e 39 (6.47%) situam-se no estado do CE.

Tabelas 5.2 - Número de poços-colunas amostrados por Unidade Federativa

UF	Quant.	%
RN	564	93,53
CE	39	6,47
Total	603	100,00

Fonte: Dados fornecidos pela Petrobras

Figura 5.1 - Número de poços-coluna amostrados por estado de localização



Fonte: Dados fornecidos pela Petrobras

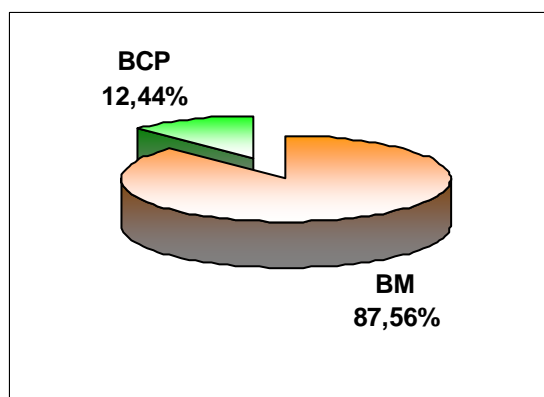
A tabela e a figura abaixo indicam a quantidade de poços funcionando de acordo com o método de elevação.

Tabelas 5.3 - Número de poço-coluna amostrado por Método de Elevação, no período de 2000 a 2006

Método de elevação	Quant.	%
BM	528	87,56
BCP	75	12,44
Total	603	100,00

Fonte: Dados fornecidos pela Petrobras

Figura 5.2 - Número de poços-coluna amostrados por métodos de elevação, no período de 2000 a 2006



Fonte: Dados fornecidos pela Petrobras

Os dados indicam que do total de poços amostrados, 87,56% produzem através do Bombeio Mecânico e apenas 12,44% através do BCP. Estes dados aproximam-se do achado na literatura, onde afirma que 80% dos poços produtores de petróleo do mundo funcionam através do Bombeio Mecânico (Costa ,2004).

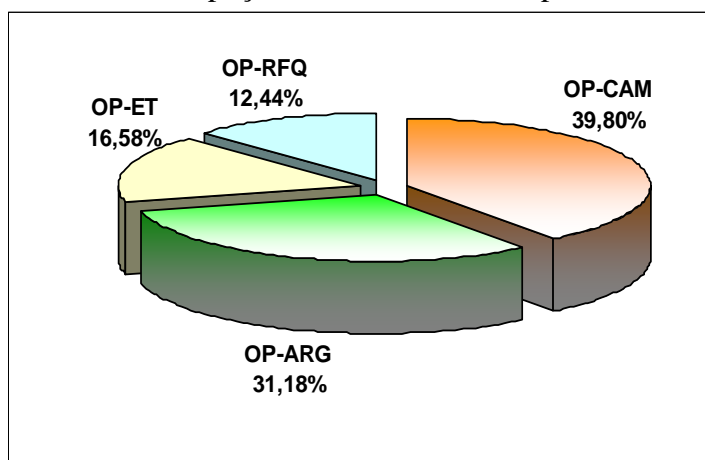
O número de poços-coluna das unidades operacionais por método de elevação é apresentado na Tabela 5.4.

Tabelas 5.4 - Número de poços-coluna amostrados por Unidade operacional

Unidade Operacional	Quant.	%	BCP	%	BM	%
OP-CAM	240	39,80	32	13,33	208	86,67
OP-ARG	188	31,18	34	18,08	154	81,92
OP-ET	100	16,58	3	3,00	97	97,00
OP-RFQ	75	12,44	6	8,00	69	92,00
Total	603	100,00	75		528	

Fonte: Dados fornecidos pela Petrobras

Figura 5.3 - Número de poços-coluna amostrados por Ativo de produção



Fonte: Dados fornecidos pela Petrobras

Nota-se que a Unidade Canto do Amaro (OP-CAM) contribuiu com maior parcela na amostra de poços, com 39.80% dos poços amostrados e a Unidade Fazenda Riacho da Forquilha (OP-RFQ) contribuiu com menor frequência (12.44%).

A Tabela 5.5 mostra interrupções por motivos mais frequentes ocorridas nos poços-coluna amostrados, no período de 2000 a 2006. Observa-se que, dentre todos os motivos relacionados a causa de interrupções, os mais frequentes são: falta de energia da Petrobras (30,10%), falta de energia elétrica – concessionária (10,81%), problema no sistema turbo-gerador (9,76%), problema na linha de alta tensão da Petrobras com 5,32% das causas e problema na unidade de bombeio (15,32%).

Tabelas 5.5 - Número de interrupções ocorridas nos poços-coluna

Motivo	Quant.	%	% acumulada
EPE - falta de energia elétrica - Petrobras	8824	30,10	30,10
ECO - falta de energia elétrica - Concessionário	3168	10,81	40,91
ETG - Sistema turbo-gerador de energia	2861	9,76	50,67
LAT - problema na linha de alta tensão Petrobras	1778	6,07	56,73
UUB - problema na unidade de bombeio	1561	5,32	62,06
Outros	11123	37,94	100,00
Total	29315	100,00	-

Fonte: Dados fornecidos pela Petrobras

A Tabela 5.6 exhibe as providencias tomada em relação às ocorrências de anormalidade dos poços. Como pode-se observar as providencias mais comuns são operações auxiliares e manutenções. A sonda de produção aparece com 12,11% das causas.

Tabelas 5.6 - Número de providencias tomadas em relação as anormalidade dos poços-colunas

Providências	Quant.	%
AO - Operações auxiliares	357	59,20
MA - Manutenção	96	15,92
SP - Sonda de produção	73	12,11
OP - Operação da instalação/poço	30	4,98
PD - Pendências	30	4,98
IN - Instalação/Construção	17	2,82
Total	603	100,00

Fonte: Dados fornecidos pela Petrobras

Do total de poços-colunas estudadas, 475 apresentaram falhas no período estudado, correspondendo a 78,77% dos casos e 128 (21,23%) apresentaram censuras, isto é, poços que não apresentaram falhas relacionadas a equipamentos de subsuperfície.

Para efeito de comparação foram realizados alguns cruzamentos entre as variáveis numéricas (produção base, BSW, RGO, idade do poço, profundidade do poço e da bomba) com as variáveis categóricas (Método de elevação e unidade operacional). As tabelas e as figuras seguintes exibem os cruzamentos citados.

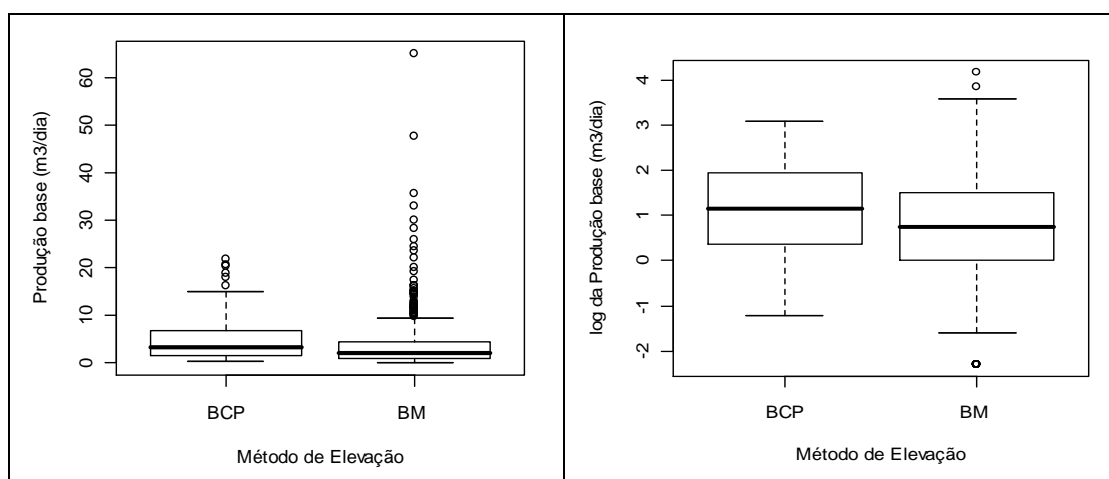
▪ **Produção base x método de elevação**

Tabelas 5.7 - Produção base por método de elevação (m³/dia)

Método de elevação	Média	1º quartil	Mediana	3º quartil	Max	Desvio-padrão
BM	3.81	1.00	2.10	4.50	65.00	5.65
BCP	5.52	1.45	3.20	6.95	21.80	5.56

Fonte: Dados fornecidos pela Petrobras

Figura 5.4 - Produção base por método de elevação



Fonte: Dados fornecidos pela Petrobras

De acordo com as medidas contidas na Tabela 5.7, nota-se que os poços BCP apresentam, em média maior produção em relação aos poços BM. A produção média para os poços produzindo através do método BM é de 3.81m³/dia com uma dispersão de 5.65 m³/dia. Para poços BCP a média é de 5.52m³/dia com dispersão de 5.56m³/dia. Estes dados indicam que os poços da Bacia Potiguar apresentam produção muito dispersa entre si. Contudo note que os poços que apresentam os maiores valores de produção no período (antes da 1º falha) usam o método BM (poços 67458, 19426, 66041 entre outros).

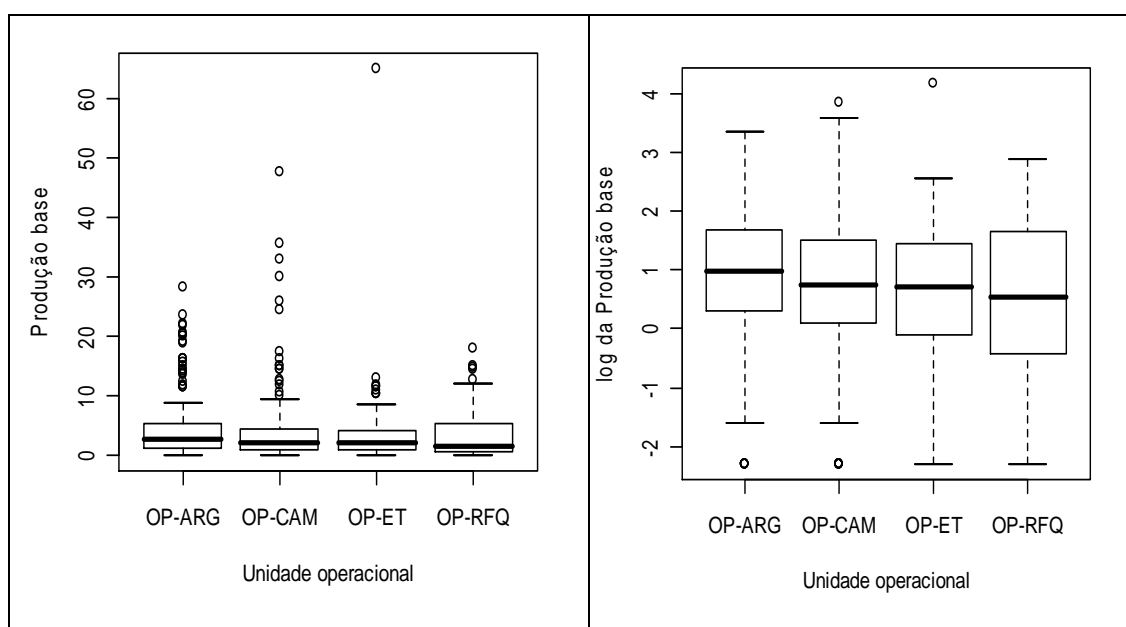
▪ **Produção base x Unidade Operacional**

Tabelas 5.8 - Produção base por unidade operacional (m³/dia)

Unidade Operacional	Média	1° quartil	Mediana	3° quartil	Max	Desvio-padrão
OP-CAM	3.81	1.10	2.10	4.50	47.60	5.73
OP-ARG	4.59	1.38	2.65	5.40	28.40	5.31
OP-ET	3.76	0.90	2.05	4.20	65.00	6.93
OP-RFQ	3.59	0.65	1.70	5.53	18.00	4.28

Fonte: Dados fornecidos pela Petrobras

Figura 5.5 - Produção base por Unidade Operacional



Fonte: Dados fornecidos pela Petrobras

Através da Figura 5.5 nota-se que os poços da OP-CAM e da OP-ARG apresentam maiores produção de óleo, em relação às demais Unidades. De acordo com a mediana exibida na Tabela 5.8, observa-se que os poços da OP-ARG apresentam maior produção de óleo (2.65m³/dia) e os poços da OP-RFQ apresentam em média menor produção (1.70m³/dia).

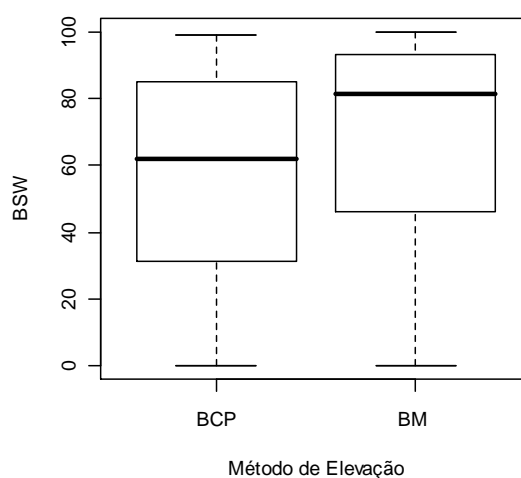
- **BSW x método de elevação**

Tabelas 5.9 - BSW por método de elevação

Método de elevação	Média	1° quartil	Mediana	3° quartil	Max	Desvio-padrão
BM	68.49	46.00	81.25	93.10	99.90	30.11
BCP	57.30	31.15	62.00	84.80	98.90	30.60

Fonte: Dados fornecidos pela Petrobras

Figura 5.6 - BSW por método de elevação



Fonte: Dados fornecidos pela Petrobras

Através da Tabela 5.9 nota-se que os poços BM apresentam maior produção de BSW, em relação aos poços BCP.

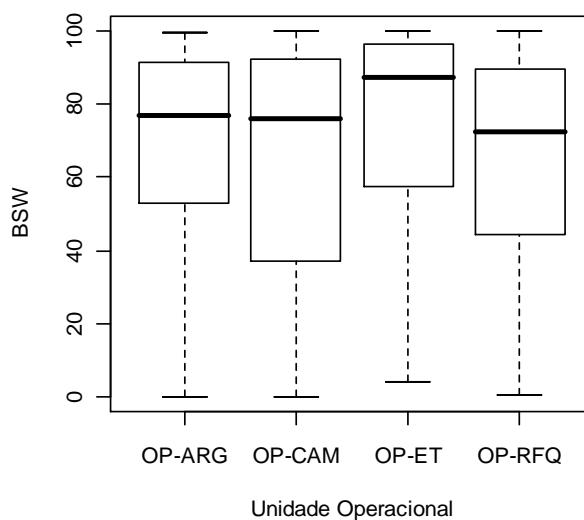
- **BSW x por unidade operacional**

Tabelas 5.10 - BSW por unidade operacional (%)

Unidade Operacional	Média	1° quartil	Mediana	3° quartil	Max	Desvio-padrão
OP-CAM	63,63	37.58	76.10	92.00	99.80	32.67
OP-ARG	67.99	53.15	77.05	91.32	99.40	29.02
OP-ET	75.23	57.77	87.20	96.33	99.80	27.32
OP-RFQ	65.11	44.45	73.20	89.55	99.90	29.32

Fonte: Dados fornecidos pela Petrobras

Figura 5.7 - BSW por unidade operacional



Fonte: Dados fornecidos pela Petrobras

A Tabela 5.10 e a Figura 5.7 exibem as medidas descritivas do BSW por unidade operacional. Observa-se que, em média a OP-ET apresenta valores mais elevados, isto significa dizer que os poços da OP-ET produzem maior quantidade de água em relação às demais unidades.

▪ **RGO x método de elevação**

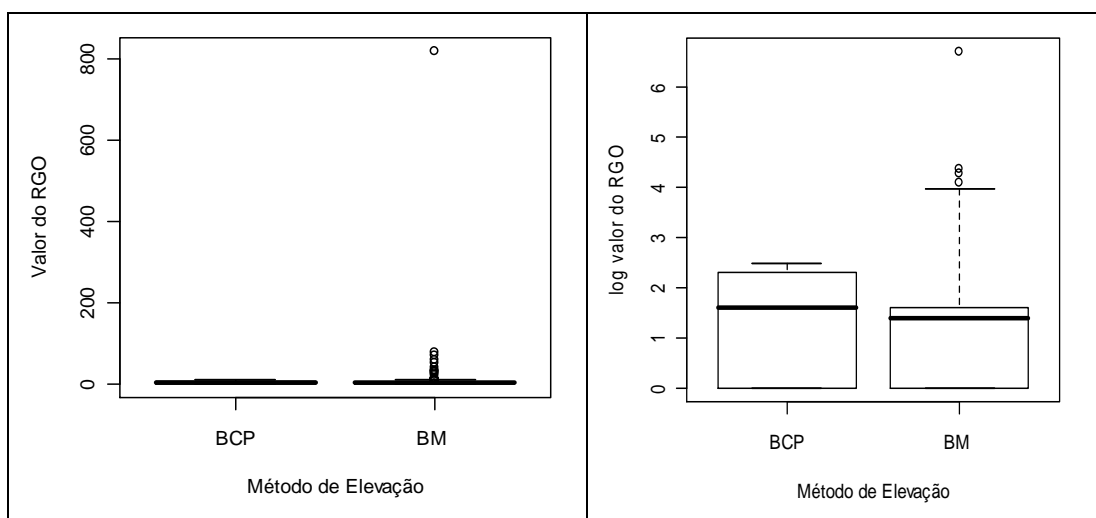
Tabelas 5.11 - RGO por método de elevação

Método de elevação	Média	1° quartil	Mediana	3° quartil	Max	Desvio-padrão
BM	6.98	1.00	4.00	5.00	816,90	36.05
BCP	5.82	1.00	5.00	10.00	11.90	4.52

Fonte: Dados fornecidos pela Petrobras

A Tabela 5.11 mostra os valores de RGO por método de elevação. Os poços BM, em média, apresentaram produção de RGO mais elevada do que os poços BCP.

Figura 5.8 - RGO por método de elevação



Fonte: Dados fornecidos pela Petrobras

▪ **RGO x por Unidade Operacional**

Tabelas 5.12 - RGO por unidade operacional

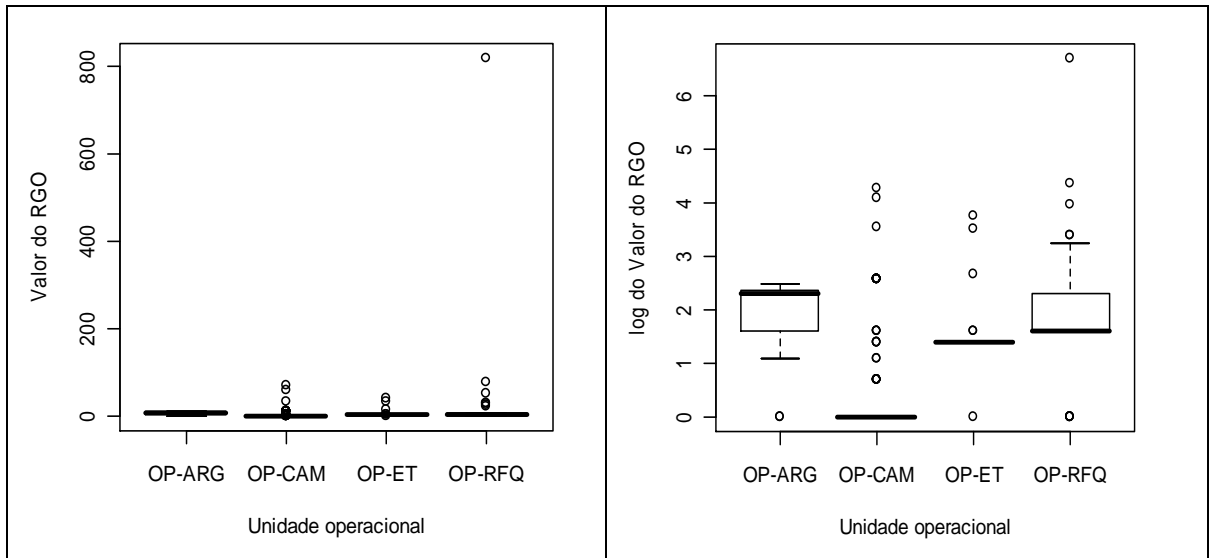
Unidade Operacional	Média	1° quartil	Mediana	3° quartil	Max	Desvio-padrão°
OP-CAM	2.37	1.00	1.00	1.00	72.00	6.75
OP-ARG	8.52	5.00	10.00	10.80	11.90	2.77
OP-ET	4.79	4.00	4.00	4.00	43.00	4.97
OP-RFQ	8.85	5.00	5.00	10.00	80.0	11.36

Fonte: Dados fornecidos pela Petrobras

Nota: As medidas dos poços da OP-RFQ foram calculadas sem a presença do valor de RGO 816,9m³

A Tabela 5.12 mostra os valores de RGO dos poços da Bacia Potiguar por unidade operacional. Nota-se que os poços da OP-ARG e da OP-RFQ apresentam maiores produção de RGO, isto é, estes poços apresentam na sua produção, maior quantidade de gás em relação a quantidade de óleo. As medidas dos poços da OP-RFQ foram calculadas sem a presença do poço 13313 que apresentava o valor do RGO muito elevado, pois ele causaria uma distorção nos resultados.

Figura 5.9 - RGO por unidade administrativa



Fonte: Dados fornecidos pela Petrobras

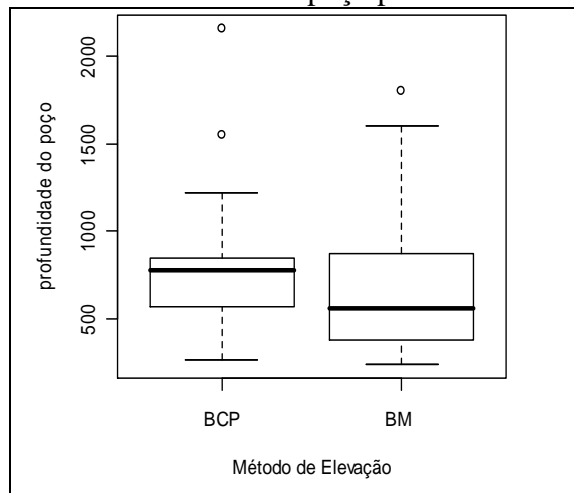
▪ **Profundidade do poço x por Método de elevação**

Tabela 5.13 - Profundidade da bomba (m) por método de elevação

Método de elevação	Média	1° quartil	Mediana	3° quartil	Max	Desvio-padrão
BM	661.6	378.0	557.0	870.8	1806.0	332.73
BCP	752.4	565.2	773.2	842.8	2160.0	268.72

Fonte: Dados fornecidos pela Petrobras

Figura 5.10 - Profundidade do poço por método de elevação



Fonte: Dados fornecidos pela Petrobras

A Tabela 5.13 mostra a profundidade do poço por método de elevação. Nota-se que os poços BCP apresentaram-se, em média, mais profundos em relação aos poços BM.

▪ **Profundidade do poço x por Unidade Operacional**

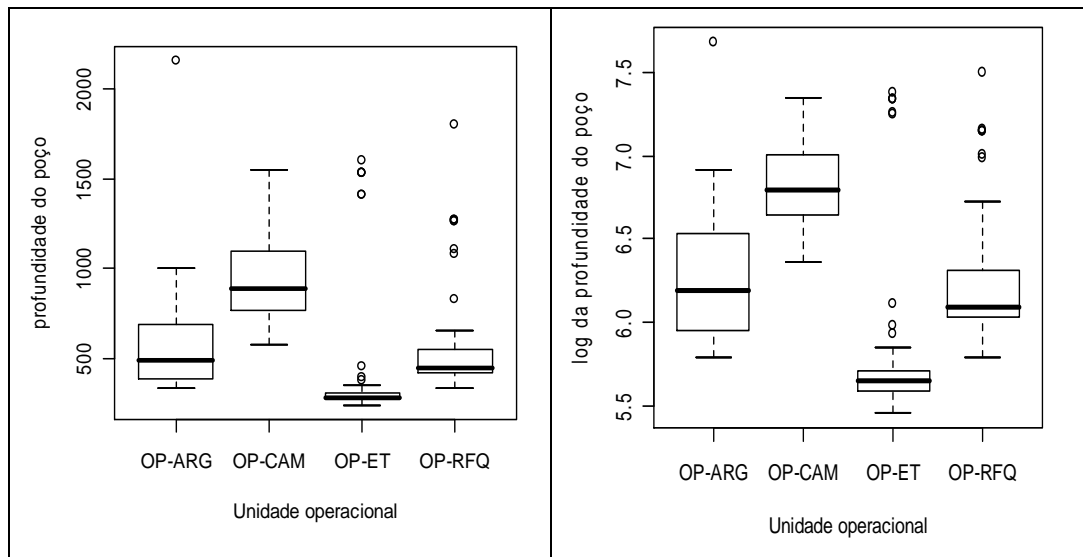
Tabelas 5.14 - Profundidade da bomba (m) por Unidade Operacional

Unidade Operacional	Média	1° quartil	Mediana	3° quartil	Max	Desvio-padrão°
OP-CAM	943.1	765.8	889.1	1098.0	1552.0	206.13
OP-ARG	547.1	368.5	490.6	682.9	2160.0	216.06
OP-ET	358.3	269.4	283.2	301.5	1604.0	293.55
OP-RFQ	543.5	418.5	442.0	551.6	1806.0	264.26

Fonte: Dados fornecidos pela Petrobras

A Tabela 5.14 mostra a profundidade do poço por unidade operacional. Nota-se que os poços da OP-CAM apresentam-se, em média, mais profundos em relação aos poços das demais OP. Já os poços menos profundos estão na OP-ET.

Figura 5.11 - Profundidade do poço por unidade administrativa



Fonte: Dados fornecidos pela Petrobras

- **Profundidade da bomba x por Método de Elevação**

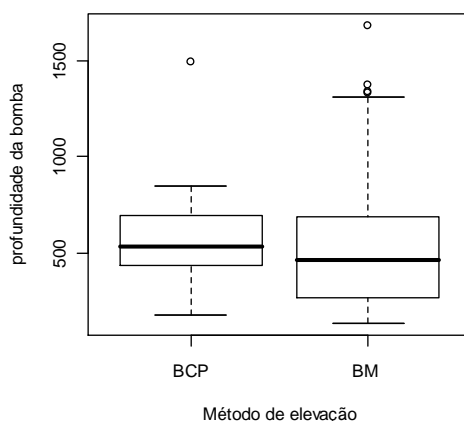
Tabelas 5.15 - Profundidade da bomba por método de elevação

Método de elevação	Média	1° quartil	Mediana	3° quartil	Max	Desvio-padrão
BM	515.6	272.8	466.5	692.2	1678.0	289.48
BCP	571.7	437.5	535.0	698.5	1493.0	187.80

Fonte: Dados fornecidos pela Petrobras

A Tabela 5.15 mostra a profundidade da bomba por método de elevação. Nota-se que os poços BCP apresentam, em média, a instalação da bomba ligeiramente mais profunda em relação aos poços BM.

Figura 5.12 - Profundidade da bomba por método de elevação



Fonte: Dados fornecidos pela Petrobras

- **Profundidade da bomba x por Unidade Operacional**

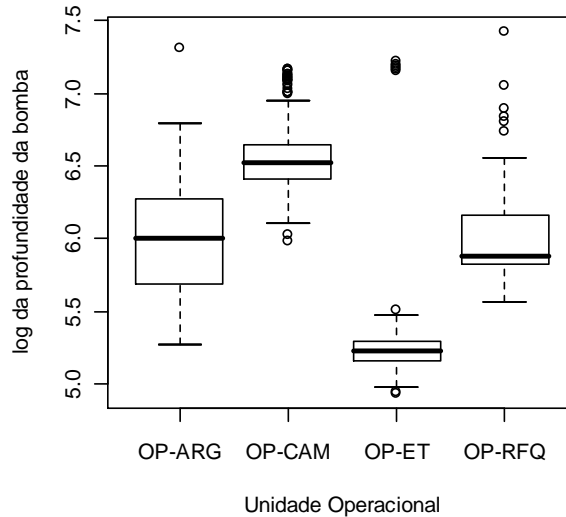
Tabela 5.16 - Profundidade da bomba por Unidade Operacional

Unidade Operacional	Média	1° quartil	Mediana	3° quartil	Max	Desvio-padrão
OP-CAM	725.6	606.8	679.5	767.0	1297.0	204.44
OP-ARG	438.8	295.8	405.5	532.8	1493.0	195.19
OP-ET	253.2	173.8	187.0	200.0	1368.0	271.59
OP-RFQ	441.9	338.0	359.0	477.5	1678.0	220.72

Fonte: Dados fornecidos pela Petrobras

A Tabela 5.16 mostra a profundidade da bomba por unidade operacional. Observa-se que os poços da OP-CAM apresentam-se, em média, com a instalação da bomba mais profunda em relação aos poços das demais OP.

Figura 5.13 - Profundidade da bomba por unidade administrativa



Fonte: Dados fornecidos pela Petrobras

A Tabela 5.17 exibe o número de falhas e censuras ocorridas por método de elevação e Unidade Operacional. Fazendo-se uma comparação do número de falhas entre os métodos de elevação por unidade operacional observa-se que tanto os poços BM como os poços BCP da OP-ARG apresentam um percentual de falha elevada. Entre os poços da OP-CAM percentual de falha é mais elevado para os poços BCP. Os poços das unidades da OP-RFQ e na OP-ET funcionando através do método BCP é muito pequeno, portanto esta comparação não é muito informativa. Apesar disso, pode-se dizer que no BM existe um percentual muito elevado de falha.

Tabela 5.17 - Número de falhas e censuras por método de elevação e unidade operacional ocorridas nos poços-colunas amostrados, no período de 2000 a 2006

Método de Elevação	Unidade Operacional			
	OP-ARG	OP-CAM	OP-RFQ	OP-ET
BM	Falha = 136 (83,31%) Censura = 18 (11,69%)	Falha = 145 (69,71%) Censura = 63 (30,29%)	Falha = 40 (57,97%) Censura = 29 (42,03%)	Falha = 86 (88,66%) Censura = 11 (11,34%)
BCP	Falha = 32 (94,69%) Censura = 2 (5,88%)	Falha = 28 (87,50%) Censura = 4 (12,50%)	Falha = 6 (100,0%) Censura = 0 (0 %)	Falha = 2 (66,67%) Censura = 1 (33,33%)
Total	188	240	75	100

Fonte: Dados fornecidos pela Petrobras

Com respeito ao tempo de funcionamento até a primeira falha dos poços foi observado um tempo médio de 14652,75 horas (aproximadamente 611 dias). A tabela 5.18 exibe as descritivas do tempo até a primeira falha do poço. Metade dos poços-coluna levam aproximadamente 485,42 dias até apresentar a primeira falha.

Tabela 5.18 - Medidas descritivas para o tempo até a falha do poço-coluna correspondente a amostra analisada no período de 2000 a 2006

Quartis	Tempo (horas)	Tempo (Dias)
25%	4412,3	184
50%	11650,0	485
75%	27827,6	1159

A Tabela 5.19 apresenta o tempo médio por covariável. Nota-se que os poços funcionando através do método BM levam em média, 773 dias para apresentar uma falha. Para os poços BCP, este tempo é aproximadamente 638 dias, aparentemente indicando que os poços BM são mais duráveis. Verifica-se também que os poços localizados na OP-CAM apresentam maior tempo de funcionamento em relação aos poços das outras Unidades. Com relação a produção, observa-se que, em média, os poços com maior produção apresentam menor tempo de funcionamento. Poços com maior produção de BSW apresentam em média maior tempo de funcionamento. Observa-se também que os poços mais profundos levam mais tempo para apresentar falha. Com relação a profundidade da bomba, nota-se que poços com instalação mais profunda apresentam maior tempo de funcionamento. Os poços mais jovens apresentam menos tempo de funcionando em relação os poços mais antigos.

Tabela 5.19 - Estimativa das medidas descritivas para o tempo até a falha do poço-coluna, por covariável, no período de 2000 a 2006

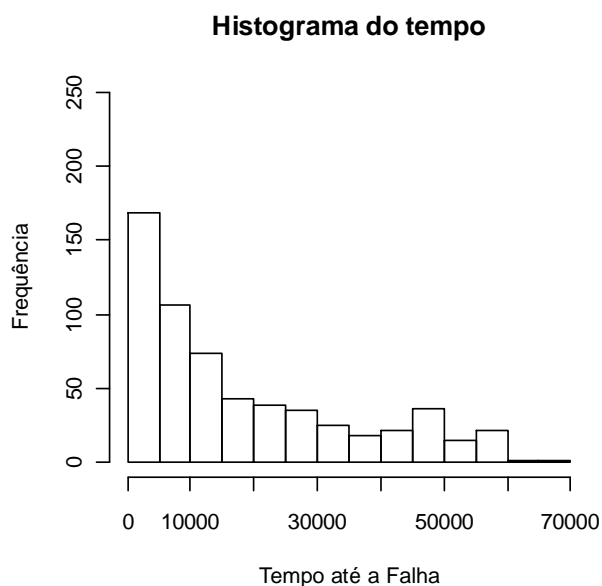
Covariável	Tempo Médio (horas)	Dias
Método de Elevação		
BCP	15314,15	638,08
BM	18543,39	772,64
Unidade Operacional		
OP-RFQ	11681,31	486,71
OP-ARG	9548,56	397,86
OP-ET	10128,15	422,02
OP-CAM	22701,06	945,90
Produção Base		
<= Med = 2,60	16905,69	704,41
> Med = 2,60	12312,79	513,03

Produção BSW		
< = Med = 77,00	13537,14	564,04
> Med = 77,00	15773,08	657,21
Profundidade do poço		
< = Med = 566,00	9920,93	413,37
> Med = 566,00	19404,55	808,52
Profundidade da bomba		
< = Med = 458,00	10057,41	419,05
> Med = 458,00	19267,49	802,81
Idade do poço (Anos)		
< = Med = 6,71	11960,46	498,35
> Med = 6,71	17357,41	723,22

5.2 - ANÁLISE NÃO-PARAMÉTRICA

A figura 5.14 apresenta a distribuição para os tempos de falha dos poços desconsiderando as covariáveis. Através do histograma do tempo nota-se uma frequência muito elevada de falha de poços no intervalo de 0 a 50000 horas. Percebe-se também uma frequência de poços muito baixa nos tempos mais elevados, apenas 41 poços apresentaram tempo de falha entre 40000 e 60000 horas.

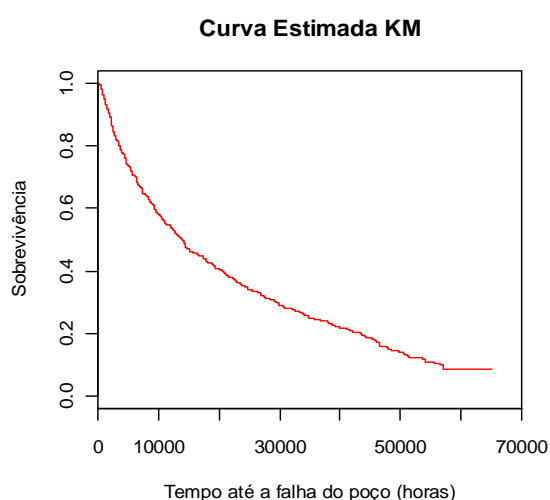
Figura 5.14 - Histograma do tempo até a primeira falha do poço



5.2.1 - ESTIMADOR KAPLAN-MEIER

Uma das análises dos dados foi construir o gráfico da função de sobrevivência estimada pelo K-M. para o tempo até a falha do poço-coluna. Este estimador é calculado sem considerar um modelo probabilístico associado aos dados. A Figura 5.15 exibe a curva de sobrevivência para o tempo até a falha.

Figura 5.15 - Curva de sobrevivência estimada a partir do estimador de Kaplan-Meier

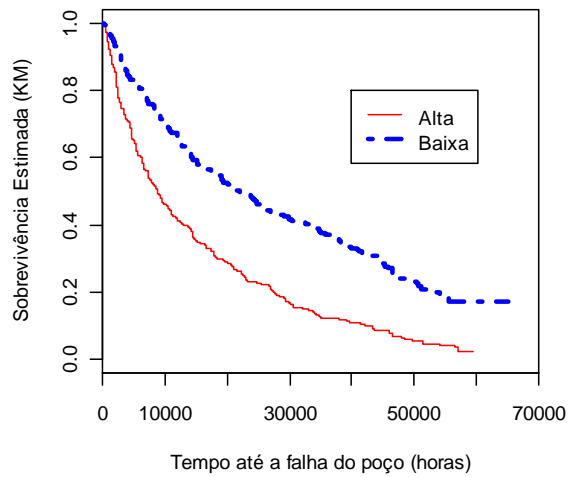


Através da figura acima é possível encontrar probabilidade de sobrevivência para diferentes tempos. Por exemplo, a probabilidade de um poço falhar em 13932 horas (equivalente 580,5 dias) é de 50%. Outra informação que se pode extrair da figura é que 40% dos poços não apresentaram falhas após 20000 horas (aproximadamente 833 dias) do início do estudo.

5.2.2 - COMPARAÇÃO ENTRE AS CURVAS DE SOBREVIVÊNCIA

As figuras a seguir apresentam as curvas de sobrevivência estimada pelo método KM (capítulo 2.3.2) para cada covariável. Estas figuras fornecem uma idéia das diferenças entre as distribuições do tempo de falha para diferentes níveis das covariáveis. A Figura 5.16 apresenta as curvas de sobrevivência estimadas para os poços em dois níveis de produção: baixa produção (produção $< 2.20\text{m}^3/\text{dia}$) e alta produção (produção $\geq 2.20\text{m}^3/\text{dia}$).

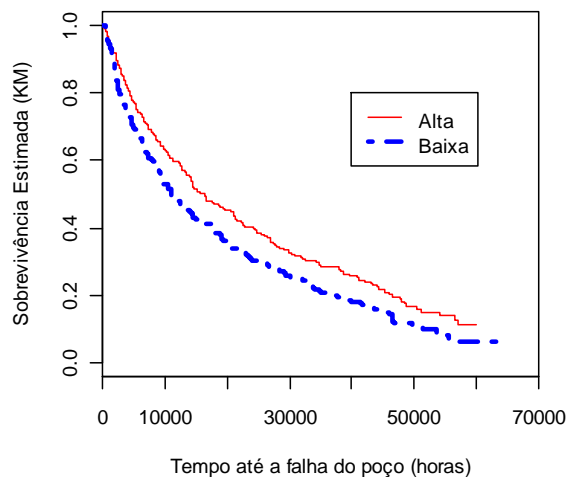
Figura 5.16 - Curva de Sobrevivência estimada pelo método Kaplan-Meier, por nível de Produção base (Baixa < 2.20 m³/dia, Alta ≥ 2.20 m³/dia)



A figura indica que os poços-coluna com baixa produção apresentam curva de sobrevivência mais elevada, quando comparada com a curva de sobrevivência dos poços com alta produção. O gráfico mostra claramente que o tempo de vida dos poços com produção elevada é menor que o tempo de vida dos poços com baixa produção

A Figura 5.17 exibe as curvas de sobrevivência em dois níveis de BSW, baixa produção do BSW (BSW < 79,40) e alta produção (BSW ≥ 79,40).

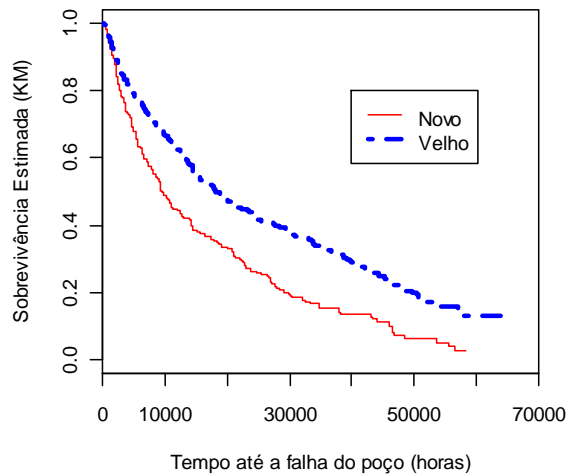
Figura 5.17 - Curva de Sobrevivência estimada pelo método Kaplan-Meier, por nível de BSW (Baixo < 79.40%, Alto ≥79.40)



A análise gráfica indica que os poços com alta produção de BSW levam mais tempo para apresentar falha em relação aos poços com baixa produção.

Através da Figura 5.18 é possível comparar as curvas de sobrevivência dos poços considerados Novos (idade < 6.71 anos) com os poços Velhos (idade \geq 6.71 anos).

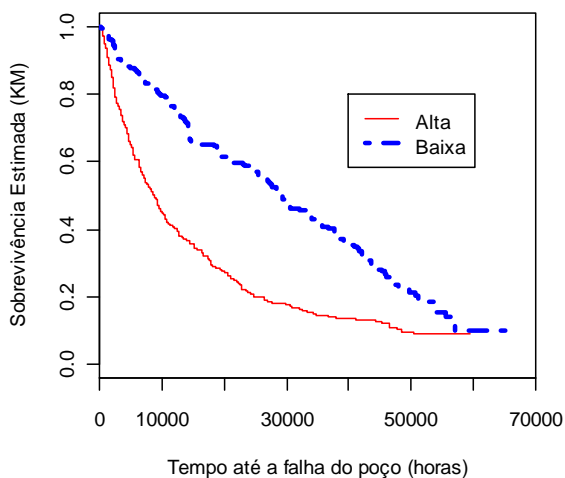
Figura 5.18 - Curva de Sobrevivência estimada pelo método Kaplan-Meier, por faixa de idade do poço (Poço Novo < 6.72 anos, Poço Velho \geq 6.72 anos)



Realizando uma análise visual do gráfico, nota-se que aparentemente a curva de sobrevivência dos poços mais antigos é superior a curva de sobrevivência dos poços mais novos. Isto indica que os poços mais velhos apresentam tempo de funcionamento mais elevado em relação aos poços mais jovens.

Através da Figura 5.19 é possível comparar as curvas de sobrevivência dos poços produzindo baixo valor de RGO ($RGO < 4 \text{ m}^3$) com a curva de sobrevivência dos poços com alta produção de RGO ($RGO \geq 4$).

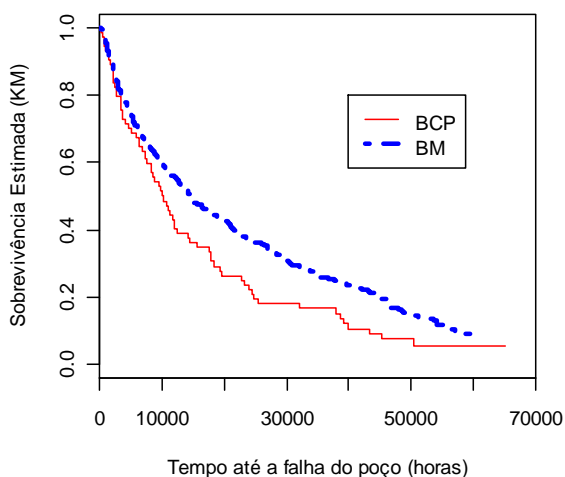
Figura 5.19 - Curva de Sobrevivência estimada pelo método Kaplan-Meier, por Nível de RGO (Baixo $< 4 \text{ m}^3$, Alto $\geq 4 \text{ m}^3$)



Como pode-se observar através da análise gráfica, o tempo de vida dos poços com baixo valor de RGO é maior em relação ao poços com alto valor de RGO.

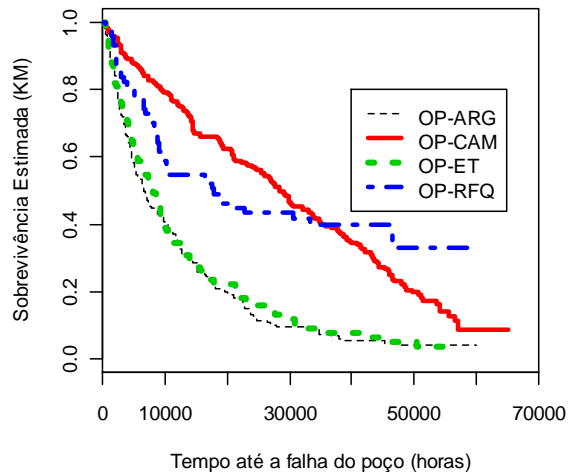
Com relação ao método de elevação dos poços, percebe-se através da Figura 5.20 uma pequena diferença nas curvas de sobrevivência. Isto significa que aparentemente os poços funcionando através do método BM apresentam curva de sobrevivência mais elevada em relação aos poços BCP.

Figura 5.20 - Curva de Sobrevivência estimada pelo método Kaplan-Meier, por método de elevação (BM e BCP)



A Figura 5.21 apresenta as curvas de sobrevivência estimadas para o tempo de falha dos poços para cada unidade operacional. Através do gráfico é possível verificar diferença entre algumas unidades operacionais.

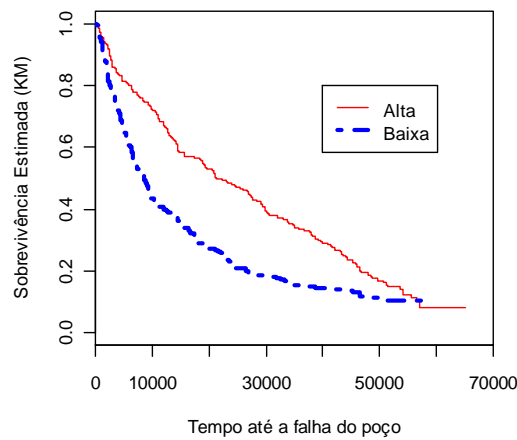
Figura 5.21 - Curva de Sobrevivência estimada pelo método Kaplan-Meier, por Unidade operacional



Nota-se que a curva de sobrevivência dos poços da OP-CAM é mais elevada em relação as demais Unidades Operacionais. Graficamente a curva de sobrevivência dos poços da OP-ARG e da OP-ET são semelhantes. O gráfico mostra que os poços da OP-RFQ apresentam probabilidade de sobrevivência inferior à probabilidade de sobrevivência dos poços da OP-CAM até aproximadamente 35.000 horas (1459 dias). Após este período, a probabilidade de sobrevivência dos poços da OP-RFQ supera a probabilidade de sobrevivência da OP-CAM. Isto parece indicar que embora os poços da OP-CAM apresentem-se mais resistente à falha no início, estes são mais afetados pelo envelhecimento que os poços da OP-RFQ. Por outro lado, a maior durabilidade dos poços da OP-RFQ após 35.000 horas pode ser pelo fato desta OP apresentar menos de 10% dos poços com BCP, que vem se mostrando menos duráveis que o BM.

Representando a profundidade da bomba tem-se a Figura 5.22 com exibição das curvas de sobrevivência para dois níveis da profundidade da bomba: Baixa profundidade (< 490m) e Alta profundidade(> = 490 m).

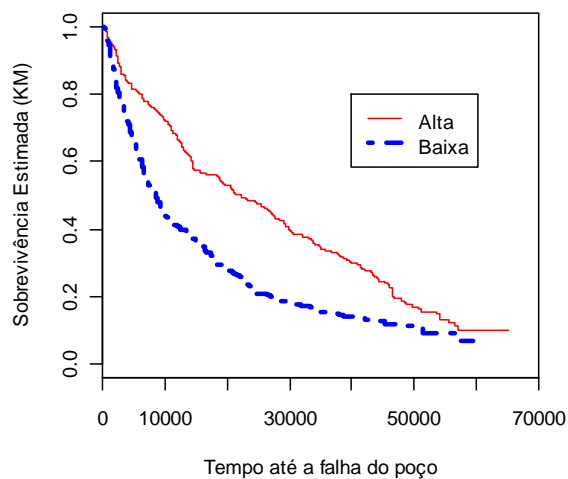
Figura 5.22 - Curva de Sobrevivência estimada pelo método de Kaplan-Meier, por nível de profundidade da bomba (baixa < 490m, alta >= 490m).



Como pode - se observar, através da análise gráfica, os poços com instalação da bomba mais profunda apresentam curva de sobrevivência mais elevada, quando comparada com a curva de sobrevivência dos poços com instalação da bomba em baixa profundidade. Isto significa dizer que os poços com bomba localizadas em maior profundidade tende apresentar maior tempo de funcionamento.

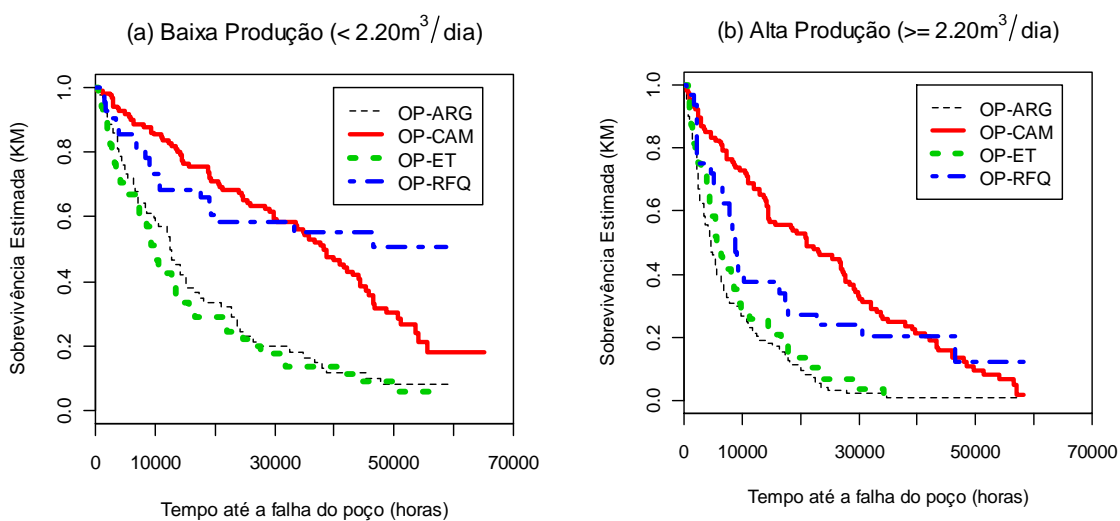
O gráfico 5.23 apresenta o tempo até a primeira falha do poço-coluna para os poços com Baixa profundidade e Alta profundidade dos poços. Nota-se que os mais profundos apresentam tempo de vida mais elevado que os poços com baixa profundidade. Isto significa dizer que os poços com maior profundidade tende apresentar maior tempo de funcionamento.

Figura 5.23 - Curva de Sobrevivência estimada pelo método de Kaplan-Meier, por nível de profundidade do poço (baixa < 675m, alta >= 675m).



Através da Figura 5.24, pode-se observar que as curvas de sobrevivência estimadas para as Unidades Operacionais OP-ARG, OP-CAM, OP-ET e OP-QFQ apresentam diferenças acentuadas, quando comparadas entre dois níveis de produção de óleo, nível alto (>=2.20 m³/dia) e nível baixo (< 2.20 m³/dia).

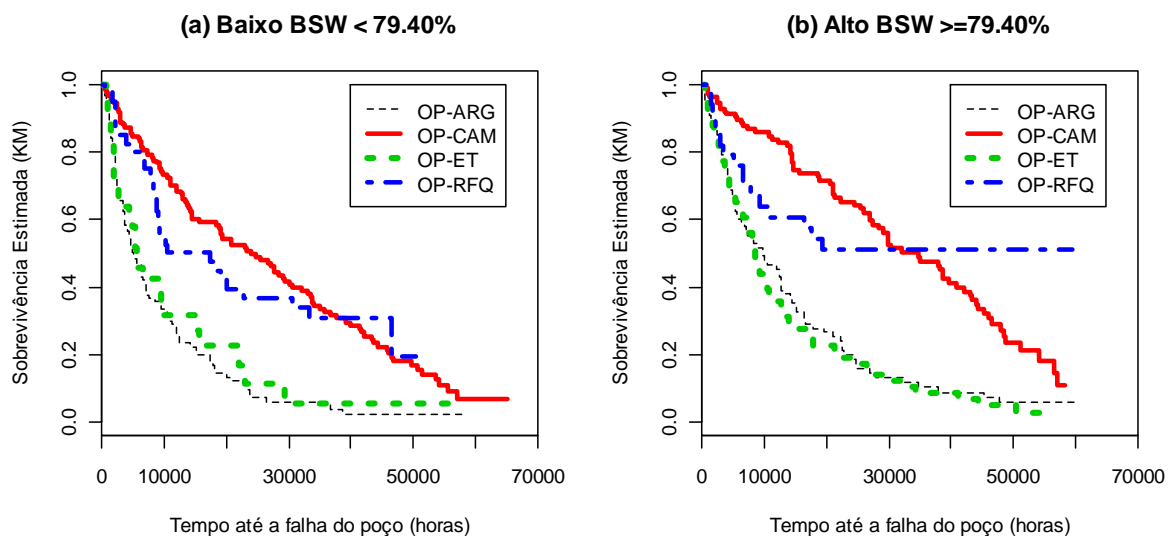
Figura 5.24 - Curva de Sobrevivência estimada pelo método de Kaplan-Meier, para cada Unidade Operacional, por nível produção de óleo (baixa < 2.20m³/dia, alta >= 2.20m³/dia).



Na OP-RFQ observa-se um decréscimo expressivo desta curva para os poços com alta produção, quando comparados aos poços com baixa produção. A figura mostra que os poços da OP-ET e OP-ARG apresentam a mesma curva de sobrevivência em ambos os níveis de produção, isto significa que aparentemente o tempo de vida dos poços destas unidades operacionais não estão sendo afetados pela produção base. A análise gráfica indica ainda, que na unidade OP-CAM o aumento da produção de óleo reduz o tempo de funcionamento dos poços. Por exemplo, para os poços com baixa produção (Figura 5.24 (a)) aproximadamente 60% deles não apresentaram falhas após 30.000 horas (1250 dias) do início do estudo. Enquanto que para os poços com alta produção (Figura 5.24 (b)), percebe-se que aproximadamente 30% dos poços não apresentaram falhas após esse mesmo tempo.

Na Figura 5.25, encontram-se as curvas de sobrevivência envolvendo as unidades operacionais e o valor do BSW.

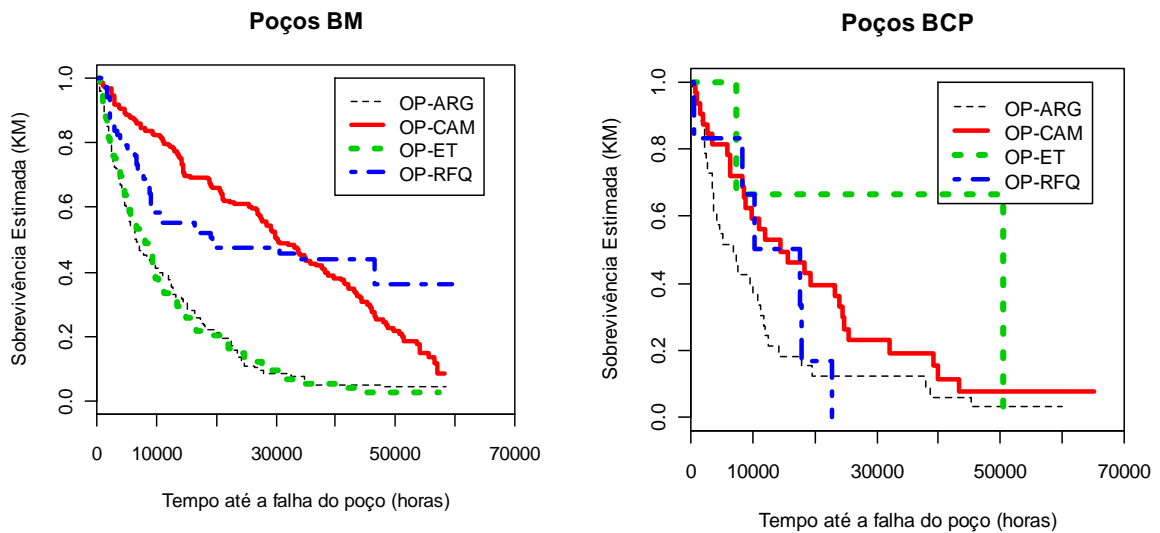
Figura 5.25 - Curvas de Sobrevivência estimada pelo método de Kaplan-Meier para cada Unidade Operacional, por nível de BSW (Baixo < 79,40%, Alto $\geq 79,40\%$)



Através das análises gráficas é possível concluir que as curvas de sobrevivências para os poços com alto valor de BSW apresentam-se mais elevada em relação as curvas dos poços com baixa produção. Visualmente é possível verificar que os poços localizados na OP-CAM apresentam probabilidade de sobrevivência mais elevada em relação aos poços das demais unidades. O gráfico mostra que a probabilidade de sobrevivência para os poços das unidades OP-ARG e OP-ET apresentam, em ambos os níveis de BSW, muito semelhantes.

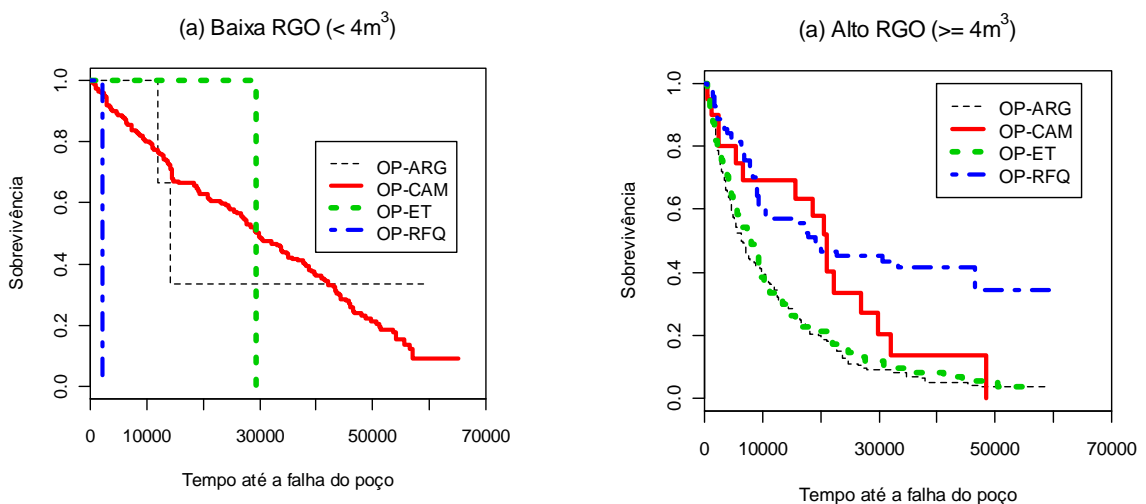
A Figura 5.26, representam as curvas de sobrevivências estimadas para os poços funcionando através do método BM e BCP, por cada Unidade Operacional.

Figura 5.26 - Curvas de Sobrevivências estimada pelo método de Kaplan-Meier para cada Unidade Operacional, por método de elevação.



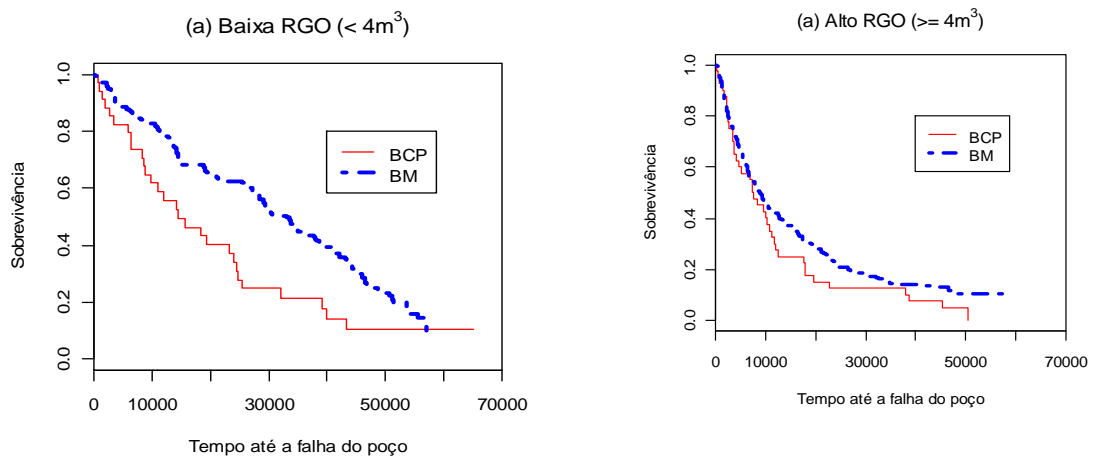
A Figura 5.26 (b) não traz muita informação pelo fato de baixo número de poços funcionando com método BCP (como visto na Tabela 5.17). Entre os poços funcionando através do BM nota-se que a Unidade da OP-CAM apresenta tempo de vida mais elevado. Além disso, o gráfico mostra que aparentemente o tempo de vida dos poços da Unidade OP-ARG e OP-ET são semelhantes.

Figura 5.27 - Curvas de Sobrevivência estimada pelo método de Kaplan-Meier para cada Unidade Operacional, em diferentes níveis por nível de RGO (Baixo $< 4 \text{ m}^3 \text{ dia}$ Alto $\geq 4 \text{ m}^3 \text{ dia}$)



A Figura 5.27 apresenta as curvas de sobrevivências estimadas por dois níveis diferentes de RGO, nível baixo ($RGO < 4.0$) e nível alto ($RGO \geq 4$), para cada Unidade Operacional. A Figura indicam que a curva de sobrevivência estimada para os poços com baixo valor de RGO, não é informativa, devido ao fato de poucos poços nesta situação (Figura 27 (a)). Com relação as curvas de sobrevivência para os poços com alto valor de RGO (Figura 27 (b)) verifica-se que os poços da OP-CAM levam mais tempo para apresentar falha.

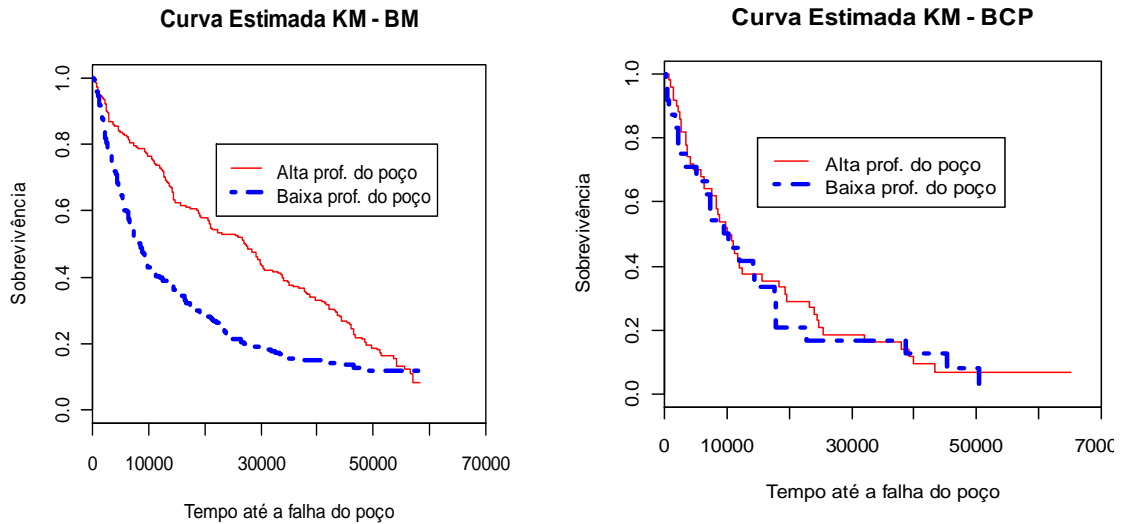
Figura 5.28 - Curvas de Sobrevivência estimada pelo método de Kaplan-Meier para cada método de elevação, por nível de RGO (Baixo $< 4 \text{ m}^3$ dia Alto $\geq 4 \text{ m}^3$ dia).



A Figura 5.28 apresenta as curvas de sobrevivências estimadas por dois níveis diferentes de RGO, nível baixo ($RGO < 4.0$) e nível alto ($RGO \geq 4$), por Método de elevação. A Figura indicam que em ambos os níveis, os poços com método de elevação BM apresentam tempo de vida mais elevado. É visto também que a diferença entre as curvas é mais acentuada entre os poços com baixo valor de RGO. Entre os poços BCP esta diferença não é muito expressiva.

A Figura 5.29 apresenta as curvas de sobrevivências estimadas para os poços em dois níveis de profundidade do poço (Alto $\geq 675,0$ m e Baixo $< 675,0$ m) pelo método de elevação. Através desta é possível comparar as curvas de sobrevivência em cada nível de profundidade, para os dois métodos de elevação, BM e BCP.

Figura 5.29 - Curvas de Sobrevivências estimada pelo método de Kaplan-Meier para cada método de elevação, por nível de profundidade do poço (Baixo < 675 m, Alto >= 675 m).



A figura mostra que para os poços BCP a probabilidade de falha em ambos os níveis é semelhante. Indicando que a durabilidade dos poços funcionando através do método BCP parece não estar sendo influenciado pela profundidade dos poços. Para os poços com método BM, nota-se que poços com alta profundidade apresentam maior tempo livre de falha, isto é, poços mais profundos e com BM levam mais tempo funcionando até apresentar a primeira falha.

5.3 - AJUSTE DO MODELO WEIBULL E DETERMINAÇÃO DOS PARÂMETROS

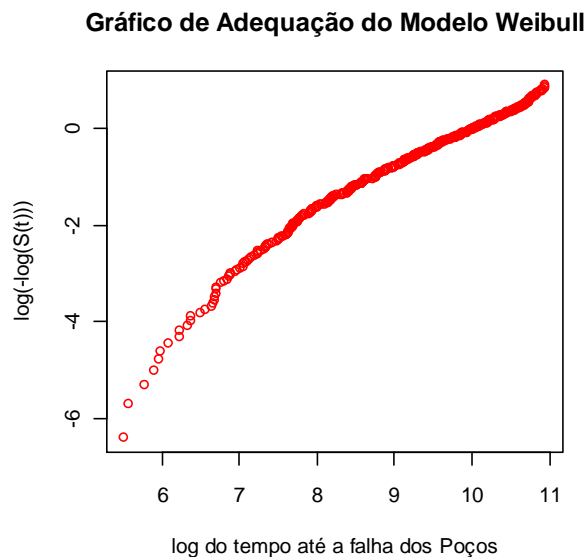
Após a análise exploratória dos dados, a etapa seguinte foi ajustar os dados a um modelo com uma determinada distribuição de probabilidade, assumindo para a variável “o tempo até a primeira falha do poço-coluna”.

A Figura 5.30 mostra o gráfico de adequação do Modelo Weibull. O critério de decisão foi baseado nos pontos do gráfico do $\log(-\log(S(t)))$ versus o \log dos tempos. O modelo será considerado adequado se os pontos se aproximarem de uma reta. Esta relação é exibida na figura abaixo que mostra um comportamento aproximadamente linear entre

$\log(-\log(S(t)))$ e o \log dos tempos, indicando um bom ajuste do modelo Weibull para os dados de tempo de falha dos poços-coluna.

Consideramos aqui o modelo Weibull, apresentado no capítulo 2, por ser amplamente utilizado para ajustar dados referentes a confiabilidade de sistemas (Accioly e Martins (1999), Bardy (2000), Frota (2003)).

Figura 5.30 - Gráfico de adequação do modelo Weibull para os dados de tempo até a primeira falha dos poços-colunas da bacia Potiguar, no período de 2000 a 2006.



Após a decisão do modelo Weibull como adequado, obtemos estimativas iniciais para os parâmetros de forma e de escala da distribuição Weibull, sem a presença de covariáveis. Utilizando o software R as estimativas obtidas foram:

$$\hat{\alpha} = \exp(\hat{\delta}) = \exp(10,025) = 22573,51 \quad \text{e} \quad \hat{\beta} = \frac{1}{\hat{\sigma}} = \frac{1}{1,14} = 0,8790246$$

Portanto os parâmetros para o modelo Weibull são $\hat{\alpha} = 22573,51$ horas e $\hat{\beta} = 0,8790246$.

As estimativas para $f(t)$ que representa aproximadamente a probabilidade que o poço falhe em um intervalo $[t, t + \Delta t)$, a curva de sobrevivência $S(t)$ que representa a

probabilidade de um poço não falhar até certo tempo t e $h(t)$ o risco de falha no tempo t , são dadas respectivamente:

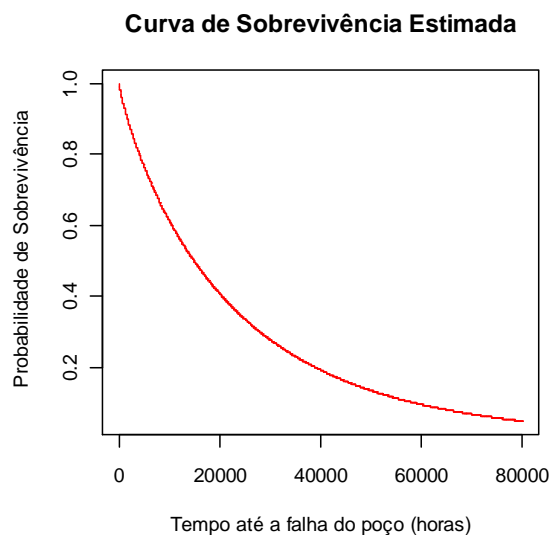
$$\hat{f}(t) = \frac{0.8790246}{22573,51^{0.8790246}} t^{0.8790246-1} \exp \left[- \left(\frac{t}{22573,51} \right)^{0.8790246} \right]$$

$$\hat{S}(t) = \exp \left(- \left(\frac{t}{22573,51} \right)^{0.8790246} \right) \quad e$$

$$h(t) = \left(\frac{0,8790246}{22573,51} \right) t^{0.8790246-1}$$

A Figura 5.31 descreve a probabilidade de um poço falhar baseado na modelo Weibull. Através desta, é obtida estimativas de interesse, por exemplo, a probabilidade de um poço falhar em um determinado instante. Por exemplo, a probabilidade de um poço funcionar até 20000 horas (aproximadamente 833 dias) sem apresentar falha é 0,5.

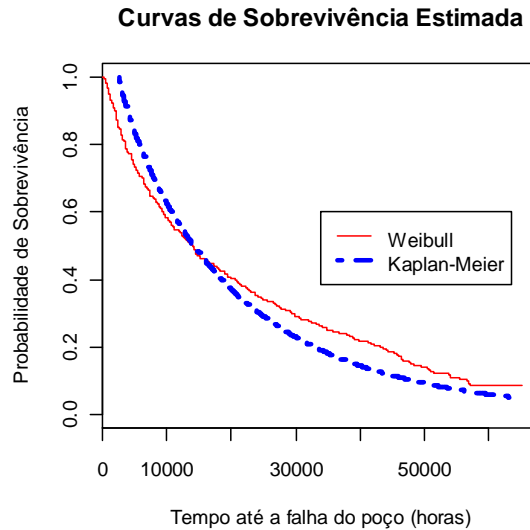
Figura 5.31 - Curva de Sobrevivência estimada baseada no modelo Weibull



A Figura 5.32 mostra conjuntamente, a curva de sobrevivência para o modelo Weibull ajustado com a curva de sobrevivência de Kaplan-Meier, pode-se observar que as

duas curvas estão próximas. Implicando que o modelo Weibull ajustado ao conjunto de dados foi escolhido adequadamente.

Figura 5.32 - Curva de Sobrevivência estimada baseada no KM e no modelo Weibull



5.3.1 - AJUSTE DO MODELO DE REGRESSÃO WEIBULL

Como o modelo Weibull foi escolhido para ajustar o conjunto de dados, iremos utilizá-lo para detectar quais covariáveis que influenciam no tempo até a primeira falha do poço-coluna. As variáveis selecionadas para compor o modelo foram: *produção base*, *valor do BSW*, *valor do RGO*, *idade do poço*, *método de elevação*, *unidade administrativa*, *quantidade de fluido previsto*, *classificação do poço*, *profundidade do poço* e *profundidade da bomba*.

O modelo considerado foi $Y = \log T_i = \beta^T X_i + \sigma W_i$ para os poços 1,2,...,n, sendo que para cada indivíduo temos:

- $t_i = \min(T_i^*, C_i)$ o tempo observado sendo $T_i = \text{tempo até a primeira falha}$ e $C_i = \text{tempo até a censura}$
- $x_i = (x_1, x_2, x_3, \dots, x_{23})^t$ as variáveis associadas

- w_i erro aleatório com distribuição valor extremo padrão e $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_{23})^t$ e σ representam os parâmetros desconhecidos que deseja-se estimar.

Utilizando-se do procedimento de seleção de covariáveis descrita na seção 2.8, elegeu-se as seguintes covariáveis significativas para explicar o tempo até a falha: Produção base, BSW, Idade do poço, RGO, Método de Elevação, Unidade Administrativa, Profundidade do poço e Profundidade da bomba. Analisou-se também possíveis interações entre algumas variáveis e constatou-se interação significativa entre unidade administrativa e método de elevação, unidade administrativa e Valor do BSW, método de elevação e profundidade do poço. Os resultados obtidos do ajuste do modelo com as covariáveis selecionadas para explicar o tempo de vida até a primeira falha do poço-coluna encontram-se na Tabela 5.20. A última coluna indica quais covariáveis foram significativas para explicar o tempo até a primeira falha.

Tabela 5.20 - Resultados do ajuste do Modelo de Regressão Weibull

Parâmetros	Variáveis Explicativas	Erro		P-valor	conclusão
		Estimativa	padrão		
β_0	(Intercepto)	7,1481	0,5289	1,30E-41	Significativo
β_1	Produção Base (X1)	-0,0454	0,0067	1,12E-11	Significativo
β_2	Valor do BSW (X2)	0,0056	0,0023	1,74E-02	Significativo
β_3	Método de elevação BM (X3)	1,7718	0,5169	6,08E-04	Significativo
β_4	Valor do RGO (X4)	-0,0028	0,0013	2,43E-02	Significativo
β_5	Idade do poço (X5)	0,0714	0,0082	4,09E-18	Significativo
β_6	Unidade Operacional OP-CAM (X6)	0,6607	0,3289	4,46E-02	Significativo
β_7	Unidade Operacional OP-ET (X7)	2,6838	0,7923	7,06E-04	Significativo
β_8	Unidade Operacional OP-RFQ (X8)	1,1049	0,4749	2,00E-02	Significativo
β_9	Profundidade do poço (X9)	0,0004	0,0007	5,23E-01	não é significativo
β_{10}	Profundidade da bomba (X10)	0,0018	0,0005	8,34E-05	Significativo
β_{11}	Método de elevação BM:Unidade Operacional OP-CAM (X11)	0,834	0,3232	9,87E-03	Significativo
β_{12}	Método de elevação BM:Unidade Operacional OP-ET (X12)	-2,1685	0,7113	2,30E-03	Significativo
β_{13}	Método de elevação BM:Unidade Operacional OP-RFQ (X13)	-0,0381	0,4694	9,35E-01	não é significativo
β_{14}	Método de elevação BM: Profundidade do poço (X14)	-0,0022	0,0007	3,00E-03	Significativo
β_{15}	Valor do BSW:Unidade Operacional OP-CAM (X15)	-0,0053	0,0032	9,71E-02	Significativo
β_{16}	Valor do BSW:Unidade Operacional OP-ET (X16)	-0,0117	0,0042	4,90E-03	Significativo
β_{17}	Valor do BSW:Unidade Operacional OP-RFQ (X17)	-0,0028	0,0048	5,64E-01	não é significativo
	Log(scale)	-0,1277	0,037	5,62E-04	

$$\hat{\sigma} = 0.88$$

Para confirmar a escolha do modelo usado no ajuste dos dados, realizou-se uma análise gráfica dos resíduos, para que o modelo escolhido possa ser considerado satisfatório o gráfico dos resíduos versus as curvas de sobrevivência dos resíduos, obtidas por Kaplan-Meier e pelo modelo exponencial, devem ser bem parecidas, Além disso, o

gráfico dos pares de pontos dos resíduos \hat{e}_i estimados pelo estimador Kaplan-Meier e pela distribuição exponencial com parâmetro ($\alpha=1$) devem ser aproximadamente linear. As Figuras 5.33 e 5.34 mostram os gráficos citados, ambos indicam que modelo Weibull parece ser aceitável,

Figura 5.33 - Curva de Sobrevivência dos resíduos obtida por Kaplan-meier e pelo método exponencial

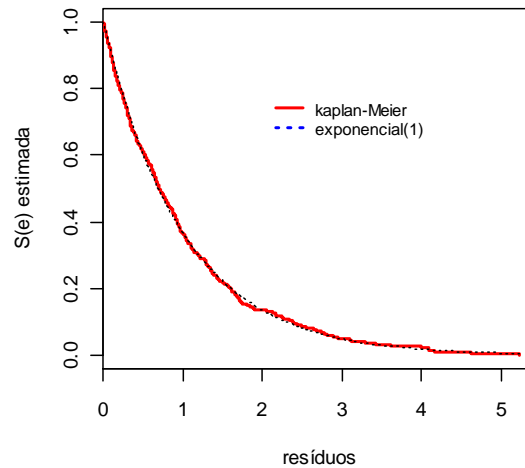
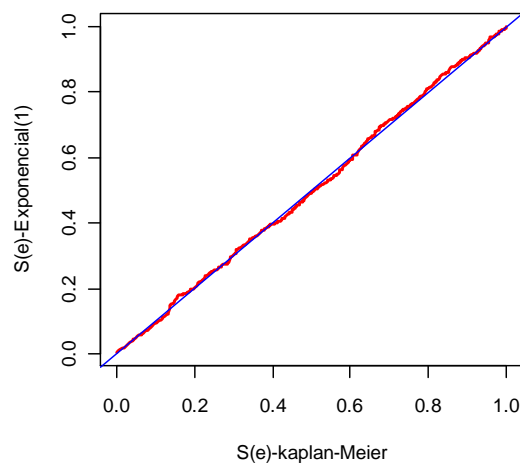


Figura 5.34 - Gráfico de pares de pontos dos resíduos obtidos por kaplan-Meier e pelo método exponencial

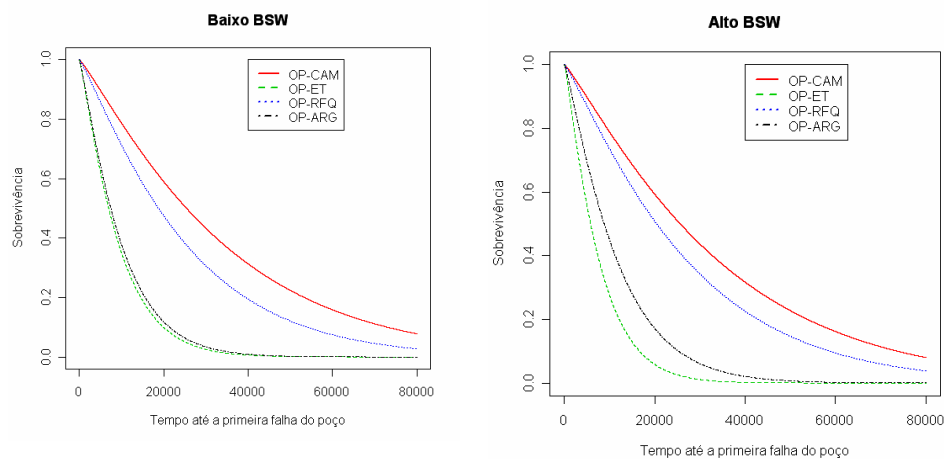


As seguintes interpretações podem ser obtidas do ajuste apresentado na Tabela 5.20.

Os poços com alta produção apresentam menor tempo de vida em relação aos poços com baixa produção ($p\text{-valor} < 0,0001$). Este resultado indica que à medida que aumenta a quantidade produzida de óleo menor a durabilidade do poço.

Com relação a variável BSW observa-se uma relação significativa positiva com o tempo de vida dos poços ($p\text{-valor} < 0,0001$), isto é, quanto maior o percentual de BSW maior é o tempo livre de falha dos poços. O modelo mostra uma interação significativa entre o valor do BSW e as unidades administrativas (Figura 3.35 ilustra esta interação). A interação indica que nas unidades OP-CAM e ARG altos valores de BSW influenciam de forma positiva na confiabilidade dos poços. O risco de falha diminui a medida que aumenta o volume de água. Este resultado mostra que poços com alta produção de água apresentam maior tempo livre de falha, indicando que a produção elevada de água não provoca dano aos equipamentos instalados na subsuperfície dos poços. Na unidade OP-CAM o aumento de BSW provoca uma redução mais acentuada no risco de falha. Para a unidade OP-ET o aumento de BSW provoca uma redução no tempo de vida dos poços. Ou seja, a produção mais elevada de água provoca maior risco de falha. Enquanto na unidade OP-RFQ a confiabilidade dos poços não parece estar sendo influenciada pelo valor do BSW.

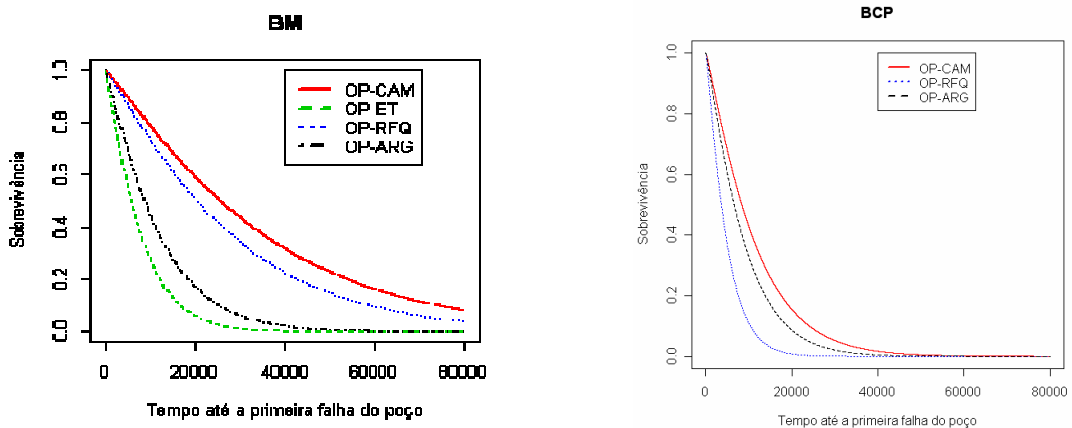
Figura 5.35 - Curvas de sobrevivência estimada através do modelo de regressão Weibull por Unidade Operacional em diferentes níveis de BSW (BSW =50 e BSW = 80)



Considerando o método de elevação dos poços verifica-se que os poços com método de elevação BM apresentam maior confiabilidade que os poços BCP ($p\text{-valor} = 0,000608$). O ajuste apontou uma interação significativa entre o método de elevação e unidade operacional (ver Figura 5.36). Os poços funcionando com método de elevação BM apresentam maior confiabilidade que os poços BCP apenas nas unidades administrativas

OP-CAM e na OP-ARG. Na unidade operacional OP-ET nota-se que os poços funcionando com método de elevação BCP são mais duráveis que os poços funcionando com método de elevação BM. Com relação aos poços da OP-RFQ parece não haver diferença significativa na confiabilidade para os dois métodos de elevação.

Figura 5.36 - Curvas de sobrevivência estimada através do modelo de regressão Weibull para poço com método de elevação BM e BCP por Unidade Operacional



Quanto maior o valor de RGO menor é o tempo livre de falha dos poços (p -valor = 0,0243). Aparentemente este resultado parece indicar que a presença de gás afeta consideravelmente o desempenho dos equipamentos de subsuperfície dos poços, levando-o a apresentarem maiores risco de falha.

Com relação a idade do poço nota-se uma relação significativa com o tempo de vida dos poços (p -valor < 0,0001). Poços mais jovens apresentam menor durabilidade em relação aos poços mais antigos. Este fato se deve a um melhor conhecimento do poço, isto é, ao longo do tempo as observações referentes às condições ambientais e características dos equipamentos instalados são incorporadas ao dia-a-dia do poço proporcionando uma melhoria no funcionamento aumentando o tempo de vida dos mesmos.

A profundidade da bomba apresentou relação com a durabilidade dos poços (p -valor < 0,0001). Verifica-se que poços com instalação da bomba mais profunda apresentam maior tempo de funcionamento em relação aos poços com a instalação da bomba em menor profundidade.

A interação entre método de elevação e profundidade do poço revelou relação significativa com o tempo de falha dos poços. Esta interação apontou que a confiabilidade dos poços BCP não está sendo afetada pela profundidade do poço, isto é, para os poços BCP o tempo de funcionamento não está relacionado com a profundidade do poço. Entre os poços BM constata-se que os poços mais profundos apresentam maiores risco de falha em relação aos poços com menor profundidade.

6. CONCLUSÃO E RECOMENDAÇÕES

Neste capítulo é apresentada uma síntese geral da dissertação, os principais pontos avaliados em cada capítulo, as conclusões e as recomendações de acordo com os resultados obtidos. Além disso, serão apresentadas as limitações e as direções para trabalhos futuros.

Este é composto por 6 partes: Pesquisa Bibliográfica, metodologia da pesquisa, limitações do trabalho, recomendações e conclusões.

6.1 - PESQUISA BIBLIOGRÁFICA

As técnicas de análise de sobrevivência aplicada neste estudo são empregadas quando pretende-se analisar o tempo até a ocorrência de um evento de interesse. Esta técnica está associada a muitas áreas, especialmente a área médica e na engenharia. Na área médica, o evento de interesse é o tempo transcorrido entre o tempo que o indivíduo entrou no estudo até o instante que apresentou características que indique o término do estudo. Este evento pode ser o tempo que o paciente levou até a morte ou o tempo até cura de uma doença, etc. Na engenharia, o evento de interesse pode estar relacionado ao tempo até a falha (degradação ou parada de funcionamento) de um determinado item. Por exemplo, o evento de interesse pode ser: O tempo até ocorrência de falha em um sistema ou tempo até a ocorrência de um reparo.

Segundo Lawless (1982) a maioria dos trabalhos estatísticos desenvolvidos referentes à área de confiabilidade foi aplicado ao tempo de vida de equipamentos e sistemas.

O conjunto de dados sobre confiabilidade é caracterizado por dois componentes: *falhas* e *censuras*, os quais conceituem a resposta. A variável de interesse em estudo de análise de sobrevivência é uma variável aleatória positiva e contínua e é usualmente dada por três funções importantes, que são: a função de densidade de probabilidade $f(t)$ que representa aproximadamente a probabilidade de que o poço venha apresentar falha dentro

de um intervalo de $[t, t + \Delta t)$, a função de sobrevivência $S(t)$ que representa a probabilidade de um poço não apresentar falha até certo tempo t e a função de risco $h(t)$ que especifica a taxa de falha instantânea no tempo t .

Antes da aplicação do ajuste do modelo paramétrico aos dados de sobrevivência deve-se sempre realizar uma análise descritiva das covariáveis através do Estimador Kaplan-Meier (Kaplan e Meier, 1958), estas medidas descrevem a curva de sobrevivência e servem para estimar a probabilidade de ocorrer falhar em um determinado instante.

Na literatura de sobrevivência existem vários modelos paramétricos que tratam de dados de vida. Alguns modelos ocupam posição de destaque por se adequar a várias situações práticas. Com isso vem sendo usados com bastante frequência, entre outros modelos, temos: modelo exponencial, weibull, log-normal, etc.

Ao determinarmos o modelo probabilístico adequado ao conjunto de dados em análise devem fazer estimativas dos parâmetros no modelo. Um dos métodos utilizados para obter tais estimativas é o método de máxima verossimilhança. O método funciona da seguinte forma: baseado nos resultados obtidos pela amostra, qual é a distribuição entre todas aquelas definidas pelos possíveis valores de seus parâmetros, com maior possibilidade de ter gerado tal amostra? Em outras palavras, se, por exemplo, a distribuição de falha é a Weibull, para cada combinação diferente de α e β têm-se diferentes distribuições de Weibull. O estimador de máxima verossimilhança escolhe aquele par de α e β que melhor explique a amostra observada (Colosimo e Giolo, 2006).

6.2 - METODOLOGIA DA PESQUISA

O estudo foi caracterizado como um estudo exploratório e explicativo. A população alvo foi os poços terrestres produtores de óleo da Bacia Potiguar sujeitos aos métodos de elevação artificial por bombeio mecânico e bombeio por cavidades progressivas. A unidade amostral considerada foram as colunas de produção instaladas em um mesmo poço. A amostragem foi definida com base na população de 4.058 poços terrestres, sendo obtida uma amostra inicial de 450 poços-coluna. Mas no período analisado vários poços-colunas entraram em operação e mudaram de coluna, conseqüentemente a amostra passou para 603 poços-colunas. Os dados foram coletados de dois sistemas de coleta de

informação da Petrobras (o SIP e o SEP). Onde as seguintes variáveis foram consideradas para análise: produção base, valor do BSW, Valor do RGO, método de elevação, idade do poço, unidade administrativa, profundidade do poço (m), profundidade da bomba (m) e classificação do poço, analisando a existência de censuras ou falha e o tempo até a falha a primeira falha do poço que representava a variável resposta.

Os métodos utilizados para alcançar o objetivo deste estudo foram considerados satisfatórios, levando em consideração os cuidados na obtenção do banco de dados, na organização e apresentação dos resultados obtidos para obter as conclusões finais.

Após a construção da base de dados para análise, foi feito, primeiramente, uma análise exploratória com a finalidade de verificar possíveis erros envolvidos na elaboração do banco de dados. Quando não foi possível identificar algumas medidas discrepantes existentes, os poços foram excluídos para não comprometer o ajuste do modelo aos dados.

A metodologia utilizada nesta pesquisa foi considerada satisfatória e não apresentou nenhuma limitação, portanto o objetivo do trabalho foi alcançado.

6.3 - RESULTADO DA PESQUISA

Considerando os resultados da análise do modelo de regressão weibull, verificou-se que o tempo até a primeira falha do poço-coluna está relacionado com algumas covariáveis analisadas. Algumas contribuem positivamente no tempo de falha, mas quando interagem com outra covariável tendem a influenciar de forma negativa no tempo de funcionamento. Os principais resultados obtidos com a regressão foram:

Produção de óleo: Poços com maior produção apresentam menor tempo de funcionamento;

Valor do RGO: Poços com maior produção de gás tendem a apresentar falha mais rápida;

Profundidade da bomba: Em geral os poços mais profundos apresentam maior tempo livre de falha. Mas considerando o método de elevação verifica-se que os mais profundos cujo método seja BM apresentam tempo de vida menor que os poços BCP. Isto é, poços mais profundos funcionam mais tempo quando o método de elevação associado seja o BM.

Método de elevação: Os poços BM em geral são mais duráveis. Observou-se que os poços BM da unidade OP-CAM são os mais duráveis. Mas na Unidade OP-ET constatou-se que o poço com método BM apresenta risco de falha mais elevada que os poços BCP.

O valor de BSW: Nota-se que a quantidade de água produzida pelo poço influencia diretamente no tempo de falha. Altos valores de BSW provocam um aumento no tempo de vida dos poços. Considerando os poços da OP-ET verifica-se que este comportamento se inverte, isto é, alto valor de BSW provocam uma redução no tempo de funcionamento do poço.

6.4 - LIMITAÇÕES DO TRABALHO

Neste estudo, a amostragem foi composta apenas com poços-colunas terrestre funcionando com método de elevação BM e BCP e produtores de óleo. Além disso, para a análise do tempo de falha até a primeira falha do poço-coluna relacionada a um equipamento de subsuperfície não foram incluídos de fatores de falha importantes como presença de areia, parafina, ambiente, corrosão, etc.

6.5 - DIREÇÕES DA PESQUISA

Este estudo apresenta alguns resultados que poderão servir de orientações para estudos futuros e tomadas de decisões. Auxiliará em um estudo envolvendo todas as falhas dos poços-coluna da Bacia potiguar, aplicando modelos de confiabilidade mais complexo e adequado ao tipo de dados. Além disso, orientará os profissionais das principais covariáveis causadores de falha nos poços-coluna da bacia potiguar, indicando previamente quais poços apresentam maior chance de falha. Consequentemente gerando a possibilidade de aumentar a vida útil e a produção dos poços, além de diminuir possivelmente gastos desnecessários com reparo.

Outro estudo a ser realizada será uma análise separada por campo/zona produtora e verificar a influência/impacto no tempo até a falha, pois a zona está associada a alguns elementos que contribuem para a falha (corrosão, incrustação, parafina, etc).

6.6 - CONCLUSÃO

Esta dissertação propôs realizar um estudo retrospectivo sobre informações relacionadas a poços terrestres da Bacia potiguar, no período de 2000 a 2006, com método de elevação BM e BCP, com a finalidade de ajustar um modelo que indicasse o risco de falha dos poços. Neste sentido, foram identificadas as covariáveis possivelmente associadas à falha e relacionadas com o tempo de vida dos poços.

O ajuste da distribuição de Weibull aos dados dos poços da Bacia potiguar mostrou-se bastante satisfatório, confirmando o recomendado pela revisão bibliográfica.

O objetivo deste trabalho foi alcançado, mas no desenvolvimento do trabalho verificou-se que existem outros aspectos a serem analisados, com previsão de serem considerados nas análises futuras.

Os resultados da análise com os dados, no período de 2000 a 2006, revelaram que:

A grande maioria dos poços da Bacia potiguar em operação eram poços com método de elevação BM. Os poços com método de elevação BM apresentaram maior tempo livre de falha em relação aos poços com método de elevação BCP. Um dos fatores que pode estar explicando esta maior durabilidade é o fato deste método apresentar maior simplicidade de operação.

Observou-se que a idade média dos poços era de aproximadamente 8 anos. A análise mostrou que 75% dos poços apresentaram idade abaixo de 13 anos e a idade máxima dos poços foi 21 anos. Constatou-se que os poços com maior tempo de vida tendem a apresentar menores riscos de falha em relação aos poços mais jovens. Este resultado pode ser indicativo de que os poços mais jovens necessitam de maior atenção no início de operação. Pois há um aprendizado maior nos poços mais antigos.

Os poços com maior produção de óleo apresentaram maiores riscos de falha em relação aos poços com menor produção.

O modelo apontou evidências para concluir que altos valores de RGO provocam redução na durabilidade dos poços. Isto significa dizer que poços com alta produção de gás apresentam maior probabilidade de falha em equipamentos de subsuperfície.

Em relação à profundidade da bomba, constatou-se que os poços com instalação da bomba em maior profundidade tendem a apresentar maior tempo livre de falha, comparando com os poços com a bomba em menor profundidade.

A quantidade do BSW esta relacionada com o tempo de vida dos poços. Esta relação comporta-se de forma diferente dentro das Unidades Operacionais. Observou-se que nas unidades OP-CAM e da OP-ARG poços com altos valores de BSW apresentaram tempo de vida mais elevado que os poços com baixo valor de BSW. Já na Unidade OP-ET esta interação apresenta-se de forma negativa, isto é, poços com alta produção apresentaram menor durabilidade que os poços com baixa produção. Na unidade OP-RFQ não foi detectada interação significativa entre o BSW e a profundidade dos poços.

Considerando as Unidades Operacionais constatou-se que o tempo de vida dos poços depende do método de elevação utilizado pelas mesmas. Na OP-CAM e na OP-ARG os poços com método de elevação BM apresentaram maior durabilidade em relação aos poços BCP. Além disso, os poços BM da OP-CAM em geral são mais duráveis que os poços BM da OP-ARG. Considerando a Unidade OP-ET os poços BCP apresentaram-se mais duráveis que os poços BM. Enquanto na OP-RFQ parece não haver diferença significativa na durabilidade dos poços para os dois métodos de elevação

Constatou-se que o risco de falha dos poços é influenciado pela profundidade do poço quando considerado conjuntamente com o método de elevação envolvido. Observou-se que para os poços BCP a profundidade não influencia a sua durabilidade. Mas considerando os poços com método BM constata-se que poços mais profundos apresentam menor tempo de funcionamento que os poços com menor profundidade.

Por fim, concluí-se que os resultados obtidos contribui como referências para trabalho futuros relacionados aos dados da Bacia e que o modelo aqui ajustado pode ser aplicado em vários outros estudos relacionados a falha de equipamentos ou de sistemas como forma de observar o comportamento da distribuições que tenha como característica o tempo como variável resposta.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALLISON, P. D. *Survival Analysis Using the SAS System. A Practical Guide*. Cary, NC: SAS Institute Inc., 1995.
- ACCIOLY, Ricardo de Mello e Silva. **Análise da duração do tempo de vida de bombas centrífugas submersas**. Dissertação de Mestrado em Engenharia de Produção. Universidade Federal do Rio de Janeiro. RJ, 1995.
- ACCIOLY, R.M.S., MARTINS, J. A. **Análise do Tempo entre as intervenções e Duração das Intervenções com Sonda**. PETROBRAS, 1999.
- BAIN, L. J. **Statistical Analysis of Reliability and Life-Testing Models: theory and Models**. Marcel Dekker, New york, 1978
- BARBETTA, P. A. **Estatística aplicada às Ciências Sociais**. 4ª ed. Florianópolis: Ed. da UFSC, 2001
- BARDY, Mariana Bahadian. **Análise de dados de falha de equipamentos de bombeio centrífugo utilizando distribuição de Weibull**. Dissertação de Mestrado em Engenharia de Produção. Universidade Federal Fluminense. RJ, 2000.
- BOLFARINE, H., RODRIGUES, J.; ACHCAR, J. A. **Análise de Sobrevivência**. 2º Escola Nacional de Modelos de Regressão, Rio de Janeiro, 1991.
- BOLFARINE, H.; Sandoval, M. Carneiro. **Introdução à Inferência Estatística**. Rio de Janeiro: Sociedade Brasileira de Matemática, 2001.
- BOLFARINE, Heleno, Bussab, Wilton. **Elementos de amostragem**. Editora Edgard Blücher. São Paulo, 2005.

- BRASIL. Regulamento - Nº 75, de 3 de maio de 2000. **Diário Oficial [da] União**, Agencia Nacional do Petróleo – ANP. Disponível em: <<http://www.google.com.br> >. Acesso em: 15 ago.2006. procedimento para Codificação de Poços perfurados com vistas à exploração ou produção de petróleo e/ou gás.
- CAVALCANTI, G. A., BEZERRA, P. R. C., SILVA, L. N. **Uma aplicação dos testes de vida acelerados para verificar o tempo de interação de paciente submetidos a cirurgia de hérnia**. 15º SINAPE. Águas de Lindóia – Sp, 2002.
- CÉSAR J. N., Aloise D. J., de Queiroz, J. W., **Análise de confiabilidade e estimação da demanda de serviços das Sondas de Produção Terrestre – SPT's na manutenção da Bacia Potiguar – RN/CE**, Dissertação de Mestrado - Universidade Federal do Rio Grande do Norte. Programa de Pós-Graduação em Sistemas e Computação. Natal - RN, 2001.
- COLLET, D. **Modelling Survival Data in Medical Research**. 1ª. Ed. London: Chapman & Hall, 1994. 347p.
- COCHRAN, W., **Sampling techniques**. New York: Wiley, 1977.
- COLOSIMO, Enrico Antônio, Giolo, Suely Ruiz. **Análise de Sobrevivência Aplicada**. Editora Edgar Blücher, São Paulo, 2006.
- COLOSIMO, E. A., **Análise de Sobrevivência Aplicada**. 46 RBRAS/9 SEAGRO – ESALQ /USP, Piracicaba, 2001.
- COSTA, R. O., **Curso de bombeio mecânico**, Relatório técnico, Petrobrás UNRNCE/ST/ELV, 2004
- COX, D. R. **Regression Models and Life Tables (With discussion)**, J.R. Stat. Soc. A., 1972

- FREIRE, Michell P. S., **Tópicos de Análise de Sobrevivência e a Aplicação a dados Referentes ao Tempo para Conclusão de um curso de Graduação.** Monografia de conclusão do curso de estatística. Universidade Federal do Rio Grande do Norte – UFRN. Natal, 2006.
- FROTA, Helder Mamede, **Desenvolvimento de método para planejamento da manutenção de poços petrolíferos em águas profundas.** Tese de mestrado em Engenharia de Reservatório e de Exploração. Universidade Estadual do norte Fluminense – UENF. Macaé, 2003.
- GARCIA, José E. L. **A completação de poços do Mar.** Apostila Petrobras, 1977.
- GIL, A. C., **Métodos de Técnicas de Pesquisa Social,** São Paulo, SP, Editora Atlas S. A., 1999.
- HORVITZ, D. G. & THOMPSON, D. J. **A generalization of sampling without replacement from a finite universe.** Journal of the American Statistical Association, v. 47, p. 663-685, 1952.
- HOFFMANN, Rodolfo. **Análise de Regressão: Uma introdução à econometria.** 3a. ed. São Paulo: HUCITEC, 1998.
- KAPLAN, E. L., MEIER, P. **Nonparametric estimation from incomplete observations.** Journal of American Statistical Association, V.53, P. 457-481, may, 1958.
- KISH, L. **Survey sampling.** Wiley, 1965.
- KLEIBAUM, D. G. **Survival Analysis: a self-learning text.** New York: Springer-Verlag, 1996
- LAWLESS, J. F., **Statistical Models and Methods for Lifetime Data.** New York: Wiley, 1982.

- LINDQVIST, B., Molnes, E., Rausand, M., **Analysis of SCSSV Performance Data**. Reliability Engineering and System Safety, 1988.
- MARTINS, J. A., ACCIOLY, R. M. S. Análise Estatística das Falhas dos Poços de Marlim. VI ETEC, 2000.
- MIURA, Kazuo et al., **Data Warehouse: Consolidação, Flexibilidade, Facilidade e Confiabilidade nos Dados para Tomada de Decisão** III SEP – II Seminário de Engenharia de Petróleo – PETROBRAS, Rio de Janeiro, 1988.
- NASCIMENTO, J. M. A., **Simulador Computacional para Poços de Petróleo com Método de Elevação Artificial por Bombeio Mecânico**. Dissertação de Mestrado do programa de pós-graduação dão curso de Engenharia Elétrica, 2005.
- NELSON, W. **Accelerated Life Testing: Statistical Models Data Analysis and Test plants**. New York: John Wiley & Sons, 1990.
- PETROBRAS, **Revista Petrobras - edição comemorativa dos 25 anos de produção terrestre de petróleo na Bacia Potiguar - janeiro de 2005**. Disponível no site da Petrobras.
- RICHARDSON, Roberto Jarry e colaboradores. **Pesquisa social: métodos e técnicas**. São Paulo: Atlas, 1985.
- R. DEVELOPMENT CORE TEAM. R: a language and environment for Statistical computing. Viena, Áustria, 2007. ISBN 3 – 900051-00-3. Disponível em <<http://www.R-project.org>> Acesso: 30/10/2007.
- SÄRNDAL, C.-E.; SWENSSON, B. & WRETMAN, J. **Model Assisted Survey Sampling**. New York: Springer-Verlag, 1992.
- SERRA, K. V. ; SANTOS, F. A. **Bombeio Mecânico, Petrobras UN-RNCE**, 1990.

- THOMAS, J. E., **Fundamentos de engenharia de Petróleo**. Rio de Janeiro: Editora Interciências, 2001.
- WEIBULL, W. A *Statistica Representation of Fatigue Failure in Solids*. Royal Institute of Techonology, 1954.

ANEXOS

ANEXO I - ASPECTOS TÉCNICOS DA AMOSTRAGEM ALEATÓRIA ESTRATIFICADA

De acordo com Bolfarine & Bussab (2005) e Cochran (1977), considere uma população U de N unidades amostrais divididas em H grupos (ou estratos) exaustivos e mutuamente excludentes. Sejam

N_h : número de unidades amostrais no estrato h;

y_{hj} : valor da variável de interesse y associado com a unidade amostral j do estrato h; para $h = 1, 2, \dots, H$ e $j = 1, 2, \dots, N_h$. Características populacionais da variável y que são de interesse são

$$\bar{y}_{hU} = \frac{1}{N_h} \sum_{j=1}^{N_h} y_{hj} : \text{média populacional de y no estrato h} \quad (4.1)$$

$$S_h^2 = \frac{1}{N_h - 1} \sum_{j=1}^{N_h} (y_{hj} - \bar{y}_{hU})^2 : \text{variância populacional de y no estrato h} \quad (4.2)$$

$$\bar{y}_U = \frac{1}{N} \sum_{h=1}^H \sum_{j=1}^{N_h} y_{hj} : \text{média populacional de} \quad (4.3)$$

$$S^2 = \frac{1}{N - 1} \sum_{h=1}^H \sum_{j=1}^{N_h} (y_{hj} - \bar{y}_U)^2 : \text{variância populacional de y.} \quad (4.4)$$

Uma amostra aleatória estratificada (AAE) de tamanho n da população U pode ser formada selecionando-se n_h unidades no estrato h ($n_h \leq N_h$) sem reposição e com iguais probabilidades. Amostras de estratos diferentes são selecionadas independentemente umas das outras. Neste esquema amostral, portanto, a probabilidade de inclusão da unidade amostral j na amostra do estrato h é

$$\pi_{hj} = \frac{n_h}{N_h} \quad (4.5)$$

para qualquer $h = 1, 2, \dots, H$ e $j = 1, 2, \dots, N_h$. A recíproca destas probabilidades de inclusão representa o peso amostral. Assim, a unidade (h, j) tem peso amostral sob AAE dado por:

$$w_{hj} = \frac{1}{\pi_{hj}} = \frac{N_h}{n_h} \quad (4.6)$$

Seja A_h o conjunto com as unidades selecionadas para a amostra do estrato h . Um possível estimador para a média populacional \bar{y}_{hU} é o estimador de Horvitz-Thompson (HORVITZ; THOMPSON,1952).

$$\bar{y}_{HT} = \frac{1}{N} \sum_{h=1}^H \sum_{j \in A_h} w_{hj} y_{hj} = \frac{1}{N} \sum_{h=1}^H \sum_{j \in A_h} \frac{1}{\pi_{hj}} y_{hj} \quad (4.7)$$

Como conseqüência das propriedades da amostragem aleatória estratificada, mostra-se através de um cálculo simples que a variância de \bar{y}_{HT} é dada por

$$V(\hat{y}_{HT}) = \frac{1}{N^2} \sum_{h=1}^H N_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{S_h^2}{n_h} \quad (4.8)$$

e um estimador não-tendencioso desta variância é

$$\hat{V}(\bar{y}_{HT}) = \frac{1}{N^2} \sum_{h=1}^H N_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{S_h^2}{n_h} \quad (4.9)$$

sendo S_h^2 a variância de y amostral de y na amostra do estrato h . Maiores detalhes deste estimador podem ser encontrados em (SÄRNDAL; SWENSSON e WRETMAN, 1992).

Para determinar o tamanho da amostra n e alocação desta amostra entre os estratos utilizou-se o seguinte procedimento:

A partir de um tamanho da amostra inicial, n_0 , aloque as observações entre os estratos proporcionalmente ao tamanho do estrato, ou seja a amostra de cada estrato será dada por

$$n_h = \frac{N_h}{N} n_0, \quad h = 1, 2, \dots, H;$$

Depois, avalie se não há estratos com amostras pequenas.

Se em 2. houver amostras pequenas, incremente tais amostras redistribuindo-se n_1 colunas adicionais, Neste caso, o tamanho da amostra será $n = n_0 + n_1$. Caso contrário mantenha $n = n_0$;

Avalie se as margens de erro amostral, calculado com o tamanho da amostra definindo em 3., para variáveis de interesse estão dentro de níveis razoáveis.

Dois aspectos importantes deste da AAE e do estimador (4.7) é o cálculo da margem do erro de amostragem e do efeito do plano amostral comparado com amostragem aleatória

simples (AAS). A margem do erro amostral de (4.7), com grau de confiança aproximadamente $1 - \alpha$, é igual a

$$Z_{1-\alpha/2} \sqrt{\hat{V}(\bar{y}_{HT})} \quad (4.10)$$

Sendo $Z_{1-\alpha/2}$ o $100(1-\alpha/2)\%$ percentil da distribuição normal. O efeito do plano amostral do estimador de Horvitz-Thompson (Kish, 1965) é definido pela razão de variâncias

$$EPA = \frac{\hat{V}_{AAE}(\bar{y}_{HT})}{\hat{V}_{AAS}(\bar{y}_{HT})} \quad (4.11)$$

em que $\hat{V}_{AAE}(\bar{y}_{HT})$ e $\hat{V}_{AAS}(\bar{y}_{HT})$ denota a variância estimada de (4.7) sob os planos AAE e AAS, respectivamente.

ANEXO II - Tabela de codificação das covariáveis incluídas no modelo

VARIÁVEL	DESCRIÇÃO	CÓDIGO	VARIÁVEIS INDICADORAS
TEMPO	Tempo em horas, do funcionamento do poço em condições normais até a primeira falha,		
CENS		0 = Censura 1= falha	
VL_PROD_BASE	Valor da Produção base do poço	m ³	X ₁
VL_BSW	Volume do BSW (Basic Sediments and Water), É a quantidade de água produzida, conjuntamente com o petróleo, O cálculo é baseado na medição da pressão hidrostática e a altura do fluido contido em um tanque e no conhecimento das densidades de água livre e da emulsão água-óleo,	%	X ₂
MEEL_CD_METODO	Método de elevação do poço	BCP BM	X ₃ 0 (Categorias de Referencias) 1
VL_RGO	Valor do RGO – Razão Gás Óleo		X ₄
IDADE DO POÇO	Idade do poço no momento da falha		X ₅
UNID_ADM	Unidade administrativa do poço-coluna	OP-ARG OP-CAM OP-ET OP-RFQ	X ₆ X ₇ X ₈ 0 0 0 (Categorias de Referencias) 1 0 0 0 1 0 0 0 1
POCO_VL_PROF_PERF	Valor da profundidade do poço	m3	X ₉
PROF	Profundidade da bomba	m	X ₁₀
MEEL_CD_METODO: UNID_ADM	Interação entre o método de elevação e a unidade operacional	BM:OP-ARG BM:OP-CAM BM: OP-ET BM: OP-RFQ	X ₁₁ X ₁₂ X ₁₃ 0 0 0 (Categorias de Referencias) 1 0 0 0 1 0 0 0 1

MEEL_CD_METODO: POCO_VL_PROF_PERF	Interação entre o método de elevação e profundidade do poço		X₁₄
VL_BSW : UNID_ADM		BSW:OP-ARG	X₁₅ X₁₆ X₁₇
		BSW:OP-CAM	0 0 0 (Categorias de Referencias)
		BSW : OP-ET	1 0 0
		BSW: OP-RFQ	0 1 0
			0 0 1

ANEXO III - Tabelas com os Resultados de cada Etapa do Ajuste do Modelo Weibull

Tabela 1. Resultados do ajuste do modelo Weibull no primeiro passo da estratégia de seleção.

Variáveis no Modelo	Covariável	Log L	-2*Log L	g.l.	ξ_{RV}	Valor-p
Nulo	-	-5239.909	10479.82	1	-	-
CBFP_VL_PROD_BASE	I	-5215.131	10430.26	1	49.55661	1.927236e-12
CBFP_VL_BSW	I	-5233.548	10467.10	1	12.72175	0.000361427
CBFP_VL_RGO	I	-5237.374	10474.75	1	5.070959	0.02432993
IDADE DO POÇO	I	-5210.382	10420.76	1	59.05408	1.532108e-14
MEEL_CD_METODO	F	-5236.894	10473.79	1	6.03089	0.01405764
UASI_CD_UNID_ADM	F	-5183.5	10367	3	112.8189	0.00000000
POCO_VL_PROF_PERF	I	-5228.662	10457.32	1	22.49481	2.107118e-06
CLPO_CD_CLASS_POCO	F	-5237.896	10475.79	6	4.025206	0.6732652
PROF	I	-5226.805	10453.61	1	26.20721	3.066747e-07

Tabela 2. Resultados do ajuste do modelo Weibull no segundo passo da estratégia de seleção.

Variáveis no Modelo		Log L	-2*Log L	g.l.	ξ_{RV}	Valor-p
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF	-	-5110.31	10220.62	10	-	-
CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF	CBFP_VL_PR OD_BASE	-5126.014	10252.03	1	31.40855	2.090586e-08
CBFP_VL_PROD_BASE + CBFP_VL_RGO+ idade.poco+ + MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF	CBFP_VL_BS W	-5112.388	10224.78	1	4.157304	0.04145534
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ MEEL_CD_METODO	idade.poco	-5143.439	10286.88	1	66.25896	4.440892e-16

+UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF							
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ idade.poco+ +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF	CBFP_VL_RGO	-5112.152	10224.30	1	3.68393	0.05493922	
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF	MEEL_CD_METODO	-5113.559	10227.12	1	6.49892	0.01079400	
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ +MEEL_CD_METODO + +POCO_VL_PROF_PERF +PROF	UASI_CD_UNID_ADM	-5150.581	10301.16	3	80.54225	0.00000000	
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ +MEEL_CD_METODO +UASI_CD_UNID_ADM+ +PROF	POCO_VL_PROF_PERF	-5114.091	10228.18	1	7.563256	0.005957027	
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF	PROF	-5115.761	10231.52	1	10.90219	0.000960504	

Tabela 3. Resultados do ajuste do modelo Weibull após o segundo passo da estratégia de seleção.(entrada das variáveis que foram excluídas no primeiro passo)

Variáveis no Modelo	Covariável	Log L	-2*Log L	g.l.	ξ_{RV}	Valor-p
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF		-5110.31	10220.62	10	-	
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ CLPO_CD_CLASS_POCO	CLPO_CD_CLASS_POCO	-5106.965	10213.92	6	6.689674	0.3505015

Tabela 4. Resultados do ajuste do modelo Weibull após o terceiro passo da estratégia de seleção
Fica o mesmo modelo do ajuste 2.

testando interação

Tabela 4. Resultados do ajuste do modelo Weibull na estratégia de seleção das interações

Variáveis no Modelo	Covariável	Log L	-2*Log L	g.l.	ξ_{RV}	Valor-p
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF		-5110.31	10220.62	10	-	-
CBFP_VL_PIROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ CBFP_VL_PROD_BASE * MEEL_CD_METODO	CBFP_VL_PRO D_BASE * MEEL_CD_ME TODO	-5110.101	10220.20	1	0.4168041	0.5185361
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ CBFP_VL_BSW * MEEL_CD_METODO	CBFP_VL_BSW * MEEL_CD_ME TODO	-5109.933	10219.87	1	0.7543059	0.3851164
FP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ CBFP_VL_RGO * MEEL_CD_METODO	CBFP_VL_RG O* MEEL_CD_ME TODO	-5110.172	10220.34	1	0.2763270	0.5991188
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ POCO_VL_PROF_PERF *MEEL_CD_METODO	POCO_VL_PRO F_PERF * MEEL_CD_ME TODO	-5110.112	10220.22	1	0.3951331	0.5296135
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ PROF *MEEL_CD_METODO	PROF * MEEL_CD_ME TODO	-5109.809	10219.62	1	1.002514	0.3167029
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO	UASI_CD_UN ID_ADM*	-5106.601	10213.20	3	7.418382	0.05969307

+UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ UASI_CD_UNID_ADM *MEEL_CD_METODO	MEEL_CD_ME TODO					
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ UASI_CD_UNID_ADM * CBFP_VL_PROD_BASE	UASI_CD_UN ID_ADM * CBFP_VL_PRO D_BASE	-5107.67	10215.34	3	5.280395	0,1523796
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ UASI_CD_UNID_ADM * CBFP_VL_BSW	UASI_CD_UN ID_ADM * CBFP_VL_BSW	-5107.548	10215.10	3	5.524669	0.1371706
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ UASI_CD_UNID_ADM * CBFP_VL_RGO	UASI_CD_UN ID_ADM * CBFP_VL_RG O	-5099.495	10198.99	3	21.63046	7.786102e-05
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ UASI_CD_UNID_ADM * POCO_VL_PROF_PERF	UASI_CD_UN ID_ADM * POCO_VL_PRO F_PERF	-5107.501	10215.00	3	5.617341	0.1317863
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF+ UASI_CD_UNID_ADM * PROF	UASI_CD_UN ID_ADM * PROF	-5106.045	10212.09	3	8.528971	0.03625552
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ +POCO_VL_PROF_PERF +PROF + Idade.poco* MEEL_CD_METODO	Idade.poco* MEEL_CD_ME TODO	-5110.308	10220.62	1	0.003461	0.9530842

Tabela 5. Resultados do ajuste do modelo Weibull na estratégia de seleção das interações

Variáveis no Modelo	Covariável	Log L	-2*Log L	g.l.	ξ_{RV}	Valor-p
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM *	-	-5093.455	10186.91			

CBFP_VL_RGO + UASI_CD_UNID_ADM *PROF						
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM * CBFP_VL_RGO + UASI_CD_UNID_ADM *PROF	UASI_CD_UNI D_ADM* MEEL_CD_MET ODO	-5097.232	10194.46	3	7.555385	0.05615232
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM *PROF	UASI_CD_UNI D_ADM * CBFP_VL_RGO	-5102.156	10204.31	3	17.40387	0.0005836476
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO	UASI_CD_UNI D_ADM *PROF	-5095.788	10191.58	3	4.666008	0.1979522

Tabela 6. Resultados do ajuste do modelo Weibull na estratégia de seleção das interações

Variáveis no Modelo	Covariável	Log L	-2*Log L	g.l.	ξ_{RV}	Valor-p
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO	-	-5095.788	10191.58			
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + CBFP_VL_PROD_BASE * MEEL_CD_METODO	CBFP_VL_PRO D_BASE * MEEL_CD_ME TODO	-5095.315	10190.63	1	0.9442127	0.3311968
CBFP_VL_PROD_BASE+CBFP_VL_BSW+	CBFP_VL_BSW	-5095.668	10191.34	1	0.2398367	0.6243241

CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + CBFP_VL_BSW * MEEL_CD_METODO	* MEEL_CD_ME TODO					
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + PROF * MEEL_CD_METODO	PROF * MEEL_CD_ME TODO	-5095.211	10190.42	1	1.154038	0.2827055
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + UASI_CD_UNID_ADM * POCO_VL_PROF_PERF	UASI_CD_UN ID_ADM * POCO_VL_PRO F_PERF	-5094.495	10188.99	3	2.584080	0.4602874
CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + CBFP_VL_RGO* MEEL_CD_METODO	CBFP_VL_RG O* MEEL_CD_ME TODO	-5085.897	10171.79	1	19.78114	8.683536e-06
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + POCO_VL_PROF_PERF * MEEL_CD_METODO	POCO_VL_PRO F_PERF * MEEL_CD_ME TODO	-5090.019	10180.04	1	11.53799	0.0006818829
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO	UASI_CD_UN ID_ADM * CBFP_VL_PRO D_BASE	-5092.411	10184.82	3	6.753497	0.08018378
CBFP_VL_PROD_BASE+CBFP_VL_BSW+	UASI_CD_UN	-5091.291	10182.58	3	8.992915	0.02938524

CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + UASI_CD_UNID_ADM * CBFP_VL_BSW	ID_ADM * CBFP_VL_BSW					
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + MEEL_CD_METODO*idade.poco	MEEL_CD_ME TODO* idade.poco	-5095.434	10190.87	1	0.7064496	0.400625

Tabela 7. Resultados do ajuste do modelo Weibull na estratégia de seleção das interações

Variáveis no Modelo	Covariável	Log L	-2*Log L	g.l.	ξ_{RV}	Valor-p
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + CBFP_VL_RGO* MEEL_CD_METODO + POCO_VL_PROF_PERF * MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_PROD_BASE + UASI_CD_UNID_ADM * CBFP_VL_BSW	-	-5071.221	10142.44			
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM * CBFP_VL_RGO + CBFP_VL_RGO* MEEL_CD_METODO + POCO_VL_PROF_PERF * MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_PROD_BASE + UASI_CD_UNID_ADM * CBFP_VL_BSW	UASI_CD_UN ID_ADM* MEEL_CD_ME TODO	-5091.433	10182.87	3	40.4231	8.666846e-09
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF	UASI_CD_UN ID_ADM * CBFP_VL_RG	-5093.654	10187.31	3	44.86448	9.887081e-10

+PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + CBFP_VL_RGO* MEEL_CD_METODO + POCO_VL_PROF_PERF * MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_PROD_BASE + UASI_CD_UNID_ADM * CBFP_VL_BSW	O					
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + POCO_VL_PROF_PERF * MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_PROD_BASE + UASI_CD_UNID_ADM * CBFP_VL_BSW	CBFP_VL_RG O* MEEL_CD_ME TODO	-5082.117	10164.23	1	21.79063	3.04082e-06
CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + CBFP_VL_RGO* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_PROD_BASE + UASI_CD_UNID_ADM * CBFP_VL_BSW	POCO_VL_PRO F_PERF * MEEL_CD_ME TODO	5077.249	10154.50	1	12.05606	0.0005162436
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + CBFP_VL_RGO* MEEL_CD_METODO + POCO_VL_PROF_PERF * MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_BSW	UASI_CD_UN ID_ADM * CBFP_VL_PRO D_BASE	-5075.71	10151.42	3	8.976733	0.02960183
CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO +UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + CBFP_VL_RGO* MEEL_CD_METODO + POCO_VL_PROF_PERF * MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_BSW	UASI_CD_UN ID_ADM * CBFP_VL_BSW	-5075.385	10150.77	3	8.326605	0.03972267

MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_PROD_BASE			
--	--	--	--

MODELO FINAL: CBFP_VL_PROD_BASE+CBFP_VL_BSW+ CBFP_VL_RGO+ idade.poco+ MEEL_CD_METODO+UASI_CD_UNID_ADM+ POCO_VL_PROF_PERF +PROF + UASI_CD_UNID_ADM* MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + CBFP_VL_RGO* MEEL_CD_METODO + POCO_VL_PROF_PERF * MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_PROD_BASE + UASI_CD_UNID_ADM * CBFP_VL_BSW

summary(ajustefinal)

Call:

```
survreg(formula = Surv(tempofalha, Indfalha) ~ CBFP_VL_PROD_BASE +
  CBFP_VL_BSW + CBFP_VL_RGO + idade.poco + MEEL_CD_METODO +
  UASI_CD_UNID_ADM + POCO_VL_PROF_PERF + PROF + UASI_CD_UNID_ADM *
  MEEL_CD_METODO + UASI_CD_UNID_ADM * CBFP_VL_RGO + CBFP_VL_RGO *
  MEEL_CD_METODO + POCO_VL_PROF_PERF * MEEL_CD_METODO + UASI_CD_UNID_ADM *
  CBFP_VL_PROD_BASE + UASI_CD_UNID_ADM * CBFP_VL_BSW, data = tab2,
  dist = "weibull", x = T)
```

(Intercept)	7.14805	0.528941	13.5139	1.30e-41
CBFP_VL_PROD_BASE	-0.04542	0.006689	-6.7903	1.12e-11
CBFP_VL_BSW	0.00555	0.002337	2.3772	1.74e-02
MEEL_CD_METODOBM	1.77176	0.516873	3.4278	6.08e-04
CBFP_VL_RGO	-0.00283	0.001256	-2.2526	2.43e-02
idade.poco	0.07144	0.008234	8.6762	4.09e-18
UASI_CD_UNID_ADMOP-CAM	0.66066	0.328874	2.0088	4.46e-02
UASI_CD_UNID_ADMOP-ET	2.68379	0.792323	3.3872	7.06e-04
UASI_CD_UNID_ADMOP-RFQ	1.10492	0.474898	2.3267	2.00e-02
POCO_VL_PROF_PERF	0.00044	0.000689	0.6385	5.23e-01
PROF	0.00184	0.000467	3.9345	8.34e-05
MEEL_CD_METODOBM :UASI_CD_UNID_ADMOP-CAM	0.83404	0.323238	2.5803	9.87e-03
MEEL_CD_METODOBM :UASI_CD_UNID_ADMOP-ET	-2.16852	0.711278	-3.0488	2.30e-03
MEEL_CD_METODOBM :UASI_CD_UNID_ADMOP-RFQ	-0.03812	0.469396	-0.0812	9.35e-01
MEEL_CD_METODOBM :POCO_VL_PROF_PERF	-0.00218	0.000736	-2.9679	3.00e-03
CBFP_VL_BSW:UASI_CD_UNID_ADMOP-CAM	-0.00531	0.003200	-1.6591	9.71e-02
CBFP_VL_BSW:UASI_CD_UNID_ADMOP-ET	-0.01170	0.004158	-2.8132	4.90e-03
CBFP_VL_BSW:UASI_CD_UNID_ADMOP-RFQ	-0.00276	0.004791	-0.5763	5.64e-01
Log(scale)	-0.12768	0.037019	-3.4492	5.62e-04
Scale=	0.88			

Weibull distribution
Loglik(model)= -5097.8 Loglik(intercept only)= -5239.9
 Chisq= 284.22 on 17 degrees of freedom, p= 0
Number of Newton-Raphson Iterations: 6
n= 603