

Modelo de regressão Weibull para estudar dados de falha de equipamentos de sub-superfície em poços petrolíferos

Maria Aldilene Dantas^{a,*}, Dione Maria Valença^b, Michell Platiny da Silva Freire^c, Pledson Guedes de Medeiros^d, Damião Nóbrega da Silva^e, Dario José Aloise^f

^aaldil_dantas@yahoo.com.br, UFRN, Brasil

^bdione@ccet.ufrn.br, UFRN, Brasil

^cmichell.freire@gmail.com, UFRN, Brasil

^dpledson@ccet.ufrn.br, UFRN, Brasil

^edamiao@ccet.ufrn.br, UFRN, Brasil

^faloisedj@hotmail.com, UFRN, Brasil

Resumo

Este trabalho apresenta um estudo de confiabilidade em dados relativos ao tempo de vida de poços petrolíferos terrestres da Petrobras, produtores de óleo na Bacia Potiguar (RN/CE). O objetivo do estudo foi, com base em um conjunto de dados sobre ocorrências de falhas, verificar a existência do relacionamento entre o tempo de vida dos poços e algumas de suas características, como método de elevação, nível de produção, BSW (*Basic Sediments and Water*), razão gás óleo (RGO), unidade operacional de origem, entre outras. Os dados foram obtidos de um estudo retrospectivo de uma amostra com 450 poços-colunas que se encontravam em funcionamento no período de 2000 a 2006, escolhida de forma a representar todos os poços da bacia RN/CE. Foi realizada uma modelagem probabilística dos dados relativos à primeira falha através do ajuste do modelo de regressão Weibull. O modelo se mostrou apropriado para ajustar os dados e foi possível identificar, através do teste da razão de verossimilhança, quais e de que forma algumas características influenciam o tempo até a falha dos poços.

Palavras-chave

Confiabilidade. Poço de petróleo. Modelo de regressão Weibull.

1. Introdução

A Bacia Potiguar é composta por 60 campos petrolíferos, localizados entre os Estados do Rio Grande do Norte e do Ceará, abrangendo aproximadamente 4.200 poços. Atualmente, a produção é de aproximadamente 78.000 barris/dia, correspondendo a 4% da produção nacional. Estes números colocam a Bacia Potiguar em primeiro lugar no Brasil, em termos de produção terrestre, e como segunda colocada entre os produtores de petróleo. Apesar de apresentar poços surgentes (com energia suficiente para a elevação dos fluidos), essa bacia possui predominantemente poços não surgentes, ou seja, que exigem instalação de um conjunto de componentes interligados para exercer pressão no fundo do poço e elevar o fluido até a superfície. As técnicas de elevação mais utilizadas entre os poços terrestres são o bombeio mecânico (BM) e o bombeio por cavidades progressivas (BCP).

Acredita-se que esses equipamentos instalados nos poços, que realizam o bombeamento dos fluidos, podem atingir uma idade produtiva de 15 anos ou mais. Contudo, ao longo do tempo, tais equipamentos apresentam falhas que causam uma parada total de produção ou provocam deficiência em seu funcionamento. Nessas situações, é necessário realizar intervenções para que o poço volte às condições normais de operar. Os serviços necessários para o restabelecimento da produção dependem de solicitações de materiais e pessoas capacitadas. Portanto, uma análise preventiva das ocorrências de falhas pode possibilitar um acompanhamento melhor dos poços que apresentam maiores riscos de ter problemas e também garantir que materiais requeridos para o restabelecimento da produção estejam disponíveis no momento da falha do poço, reduzindo assim o tempo de manutenção.

Na literatura existem diversos trabalhos utilizando técnicas de análise de sobrevivência ou confiabilidade (LAWLESS, 1982; NELSON, 1990) para a análise da durabilidade de determinados equipamentos instalados em poços de petróleo e alguns poucos que consideram o estudo da durabilidade do poço como um sistema. Bardy (2000) apresenta uma análise de tempos de operação de equipamentos de bombeio centrífugo submerso, instalados em poços de petróleo, utilizando o modelo probabilístico Weibull (COLOSIMO; GILOLO, 2006). Esse mesmo modelo foi utilizado por Accioly e Martins (1999) na análise das intervenções no campo de Marlim e Albacora, na Bacia de Campos (Brasil). Esses autores observaram o tempo entre as intervenções e o tempo de produção dos poços que não sofreram intervenção. Linqvist et al. (1988) usam o modelo de riscos proporcionais de Cox (1972) em dados relativos a falhas das principais válvulas de segurança de sub-superfície utilizadas em 26 campos de óleo e gás no mar do Norte. Frota (2003) utilizou técnicas usuais de confiabilidade para analisar dados relativos a um período de 12 anos de intervenções em poços marítimos, da Bacia de Campos, com a finalidade de identificar as causas das intervenções e modelar através da distribuição Weibull o tempo entre falhas. Esse autor assume que após o reparo os equipamentos ficam como novos, sendo possível assim assumir independência entre os tempos. Percy e Alkali (2005, 2007) propõem modelos baseados em processos de Poisson não homogêneo para modelar o tempo entre falhas em sistemas reparáveis, e argumentam que apenas o tempo até a primeira falha permite o uso de modelos de confiabilidade usuais.

O presente estudo tem como principal objetivo a análise de uma amostra de dados sobre ocorrências de falhas em poços da Petrobras, através do ajuste de um modelo de regressão Weibull, para verificar a existência do relacionamento entre o tempo até a primeira falha e algumas características dos poços, buscando assim uma posterior atuação preventiva naqueles que apresentam maiores risco de falha.

Este trabalho está organizado em cinco seções. A Seção 2 introduz conceitos de análise de sobrevivência e detalha o modelo utilizado. A Seção 3 apresenta descrição dos dados e dos procedimentos utilizados para o dimensionamento e seleção da amostra. A Seção 4 fornece o ajuste do modelo aos dados e finalmente na Seção 5 são encontradas as conclusões do trabalho.

2. Modelo de regressão Weibull

Para cada item i , com $i = 1, \dots, n$. Seja T_i^* uma variável aleatória positiva e contínua representando o tempo até a ocorrência de uma falha, com função de densidade $f(t)$ e função de sobrevivência $S(t) = P(T_i^* > t)$, e seja C_i variável aleatória representando o tempo até a censura, ou seja, uma informação incompleta do tempo de acompanhamento. Os tempos observados são $T_i = \min(T_i^*, C_i)$ e definimos um indicador de censura (falha) δ_i que toma o valor 1 se ocorre falha e 0 se ocorre censura. Na maioria dos estudos de tempo de vida há características associadas à sobrevivência dos itens que são representadas através de quantidades chamadas variáveis explanatórias ou covariáveis. Por exemplo, o tempo até a ocorrência de falha dos poços pode depender da profundidade da bomba ou do método de elevação. Assim, associado a cada item considere um vetor de variáveis explanatórias $x_i = x_{i1}, \dots, x_{ip}$. Se tivermos interesse em determinar a relação entre o tempo até ocorrência de falha e o vetor de variáveis explanatórias, fazemos isso através do uso de um modelo de regressão (Lawless, 1982). O modelo utilizado neste trabalho foi o modelo de regressão Weibull, representado da seguinte forma (Equação 1):

$$\log(T_i^*) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \sigma v_i \quad (1)$$

em que $\beta = (\beta_0, \beta_1, \dots, \beta_p)$ e σ são parâmetros desconhecidos, $x' = (x_{i1}, x_{i2}, \dots, x_{ip})$ é o vetor de variáveis explanatórias e o erro v tem distribuição valor extremo padrão, com densidade $f(v) = \exp\{v - e^v\}$ para $v \in \mathbb{R}$. Dessa forma, T^* tem uma distribuição Weibull e, consequentemente, $\log(T^*)$ tem uma distribuição valor extremo com parâmetro de escala σ e locação $\mu(x) = x'\beta$ dependendo das covariáveis.

A estimativa desses parâmetros é obtida a partir do método da máxima verossimilhança (BOLFARINE; SANDOVAL, 2001), que consiste em resolver o sistema de equações gerado a partir da diferenciação do logaritmo da função de verossimilhança. A função de verossimilhança para este modelo é apresentada abaixo (Equação 2):

$$L(\beta, \sigma) = \prod_{i=1}^n \left(\frac{1}{\sigma} \exp\left\{ \frac{y - x'\beta}{\sigma} \right\} - \exp\left\{ \frac{y - x'\beta}{\sigma} \right\} \right)^{\delta_i} \left(\exp\left\{ -\exp\left\{ \frac{y - x'\beta}{\sigma} \right\} \right\} \right)^{1 - \delta_i} \quad (2)$$

As equações resultantes do processo de derivação geralmente não apresentam uma solução analítica, portanto, devem ser resolvidas a partir de um método numérico. O método usualmente utilizado é de Newton-Raphson (BOLFARINE; SANDOVAL, 2001).

Admitindo que o modelo de regressão Weibull seja apropriado para o estudo, dá-se início à escolha de um conjunto de variáveis regressoras que melhor explique a resposta (o tempo de vida). A seleção das variáveis explicativas pode ser baseada na comparação de diferentes modelos contendo covariáveis envolvidas na análise, a fim de selecionar aquele que melhor se ajuste aos dados. Essa comparação é realizada através do teste de Razão de Verossimilhança (BOLFARINE; SANDOVAL, 2001), com base num procedimento de seleção não automático para variáveis explicativas apresentado em Collet (1994).

3. Descrição dos dados e seleção da amostra

Foram coletados dados sobre falhas de poços de petróleo no período de janeiro de 2000 a dezembro de 2006. Os dados levantados incluem informações referentes apenas aos poços terrestres e produtores de óleo, que utilizam método de elevação BM ou BCP.

Consideramos como unidade de medida o sistema poço-coluna, visto que um único poço pode produzir através de duas colunas independentes. Especificamente assumimos que um poço-coluna funciona como 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 dos equipamentos de sub-superfície.

A variável resposta é o tempo de funcionamento do poço dentro da sua normalidade até apresentar a primeira falha, decorrente de alguma anormalidade relacionada a equipamentos de sub-superfície que tenham ocasionado uma parada total no funcionamento do poço.

Para garantir que a amostra a ser retirada fornecesse informações para todos os poços da Bacia Potiguar, optou-se por realizar a técnica de amostragem aleatória estratificada (AAE). As variáveis, profundidade da bomba, método de elevação e os dois ativos de produção (ATP-ARG – Ativo de Produção Alto do Rodrigues e ATP-MO – Ativo de Produção Mossoró) foram escolhidos para definir estratos. Para formar estratos mais homogêneos, subdividiu-se cada estrato de acordo

com faixas de profundidade, sendo duas em cada grupo com método de elevação BCP e três faixas em cada grupo com BM, formando assim um total de dez estratos.

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

- A partir de um tamanho da amostra inicial, $n_o = 400$, alocaram-se as observações entre os estratos proporcionalmente ao tamanho do estrato, isto é, a amostra de cada estrato n_h foi dada por (Equação 3)

$$n_h = \frac{N_h}{N} n_o, \quad h = 1, 2, \dots, H; \quad (3)$$

onde N_h corresponde ao número de poço-coluna dentro de cada estrato.

- Depois, percebendo-se a ocorrência de amostras pequenas em alguns estratos, foram acrescentadas 50 unidades amostrais, redistribuídas entre os estratos. Nesse caso, o tamanho da amostra ficou 450.
- Avaliaram-se as margens de erro amostral, calculadas com o tamanho da amostra definido ($n = 450$) para variáveis de interesse, estando todas dentro de níveis razoáveis.
- Um dos aspectos importantes da AAE é o cálculo da margem do erro de amostragem e do efeito do plano amostral comparado com amostragem aleatória simples (AAS). O efeito do plano amostral do estimador é definido pela razão de variâncias (Equação 4)

$$EPA = \frac{\check{V}_{AAE}(\bar{y}_{HT})}{\check{V}_{AAS}(\bar{y}_{HT})} \quad (4)$$

em que $\check{V}_{AAE}(\bar{y}_{HT})$ e $\check{V}_{AAS}(\bar{y}_{HT})$ denotam a variância estimada do estimador Horvitz-Thompson (HORVITZ; THOMPSON, 1952) sob os planos AAE e AAS, respectivamente.

- O procedimento de amostragem selecionou amostras de cada estrato aleatoriamente, sendo atribuídas iguais probabilidades de seleção para cada poço-coluna pertencente a esse estrato.

Assim, no sentido de estimar a profundidade média da bomba na população com o estimador de Horvitz-Thompson pode-se mostrar que, com um grau de confiança de 95%, a amostra retirada produziu uma margem sobre o erro amostral de apenas 8 m, o que representa um valor pequeno, considerando que a profundidade média fica em torno de 500 m.

Após a seleção foi construído um banco de dados com todas as informações sobre intervenção dos poços selecionados no período de 2000 a 2006.

4. Aplicação

Utilizando-se do procedimento de seleção de covariáveis descrito anteriormente, as seguintes covariáveis foram eleitas como significativas para explicar o tempo até a falha: produção de óleo (m^3/dia), % do BSW, valor do RGO, idade do poço, método de elevação (bombeio mecânico – BM ou bombeio por cavidade progressiva – BCP), unidade administrativa (OP-CAM – Unidade Operacional Canto do Amaro, OP-ARG – Unidade Operacional Alto do Rodrigues, OP-ET – Unidade Operacional Estreito e OP-RFQ – Unidade Operacional Riacho Forquilha), profundidade do poço (m), profundidade da bomba (m). Com base nesse banco de dados e com as datas de ocorrências de falha foi possível calcular o tempo de vida dos poços.

As figuras a seguir apresentam as curvas de sobrevivência estimadas pelo método Kaplan-Meier (KAPLAN; MEIER, 1958) para cada covariável. Estas figuras fornecem uma idéia das diferenças entre as distribuições do tempo de falha para diferentes níveis das covariáveis. A Figura 1 apresenta as curvas de sobrevivência estimadas para os poços em dois níveis de produção: baixa produção (produção $< 2,20m^3/dia$) e alta produção (produção $\geq 2,20 m^3/dia$) e 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, ou seja, o tempo de vida dos poços com produção elevada parece ser significativo menor que o tempo de vida dos poços com baixa produção.

A Figura 2 exibe as curvas de sobrevivência em dois níveis de BSW, baixa produção do BSW ($BSW < 79,40$) e alta produção ($BSW \geq 79,40$). 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 3 é possível comparar as curvas de sobrevivência dos poços considerados novos (idade $< 6,71$ anos) com os poços velhos (idade $\geq 6,71$ anos). Realizando uma análise visual do gráfico, nota-se que aparentemente a curva de sobrevivência dos poços mais antigos é superior à curva de sobrevivência dos poços mais novos. Isso indica que os poços mais antigos apresentam tempo de funcionamento mais elevado que os poços mais jovens, o que, segundo técnicos da Petrobras, ocorre em função de se ter um maior conhecimento dos poços mais velhos.

Com relação ao método de elevação dos poços, percebe-se através da Figura 4 uma pequena diferença nas curvas de sobrevivência. Isto significa que aparentemente os poços funcionando através do método BM apresentam sobrevivência mais elevada em relação aos poços BCP.

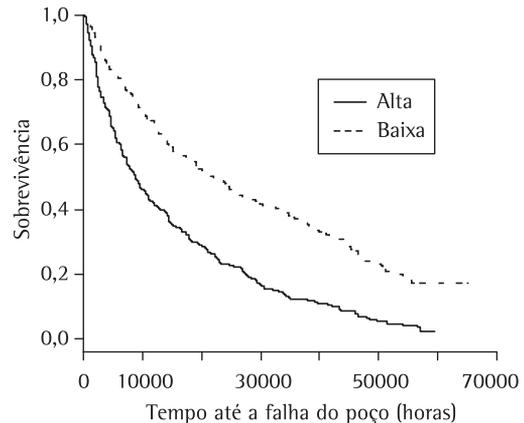


Figura 1. Curva de sobrevivência estimada pelo método Kaplan-Meier, por nível de produção base (baixa $< 2,20 m^3/dia$, alta $\geq 2,20 m^3/dia$).

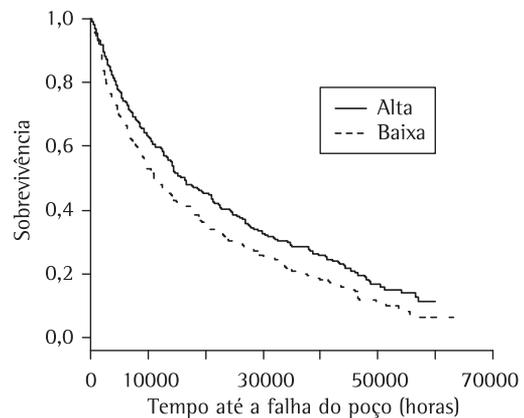


Figura 2. Curva de sobrevivência estimada pelo método Kaplan-Meier, por nível de BSW (baixo $< 79,40\%$, alto $\geq 79,40\%$).

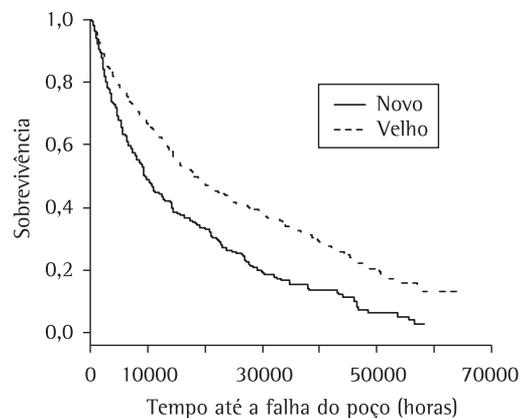


Figura 3. Curva de sobrevivência estimada pelo método Kaplan-Meier, por faixa de idade do poço (poço novo $< 6,72$ anos, poço velho $\geq 6,72$ anos).

Após a análise exploratória dos dados, foi ajustado um modelo de regressão Weibull, assumindo “o tempo até a primeira falha do poço-coluna” como a variável de interesse. Os resultados obtidos a partir do ajuste do modelo de regressão Weibull descrito pela Equação 1 utilizando o software computacional R, versão 2.6, estão apresentados na Tabela 1.

As covariáveis selecionadas para explicar o tempo até a primeira falha do poço-coluna podem ser escritas da seguinte forma (Equação 5):

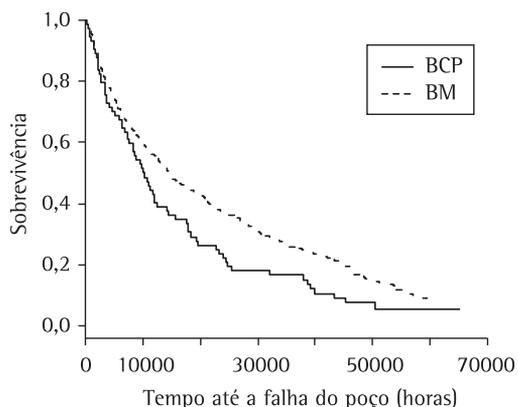


Figura 4. Curva de sobrevivência estimada pelo método Kaplan-Meier, por método de elevação (BM e BCP).

$$\begin{aligned} \hat{Y} = & 7,148 - 0,045PROD + 0,006BSW + 1,772BM - \\ & - 0,003RGO + 0,071IDADE + 0,661OP.CAM + \\ & + 2,684OP.ET + 1,105OP.RFQ + \\ & + 0,0004PROF.POCO + 0,002PROF.BOMBA + \\ & + 0,834BM * OP.CAM - 2,168BM * OP.ET - \\ & - 0,038BM * OP.RFQ - 0,002BM * PROF.POCO - \\ & - 0,005BSW * OP.CAM - 0,012BSW * OP.ET - \\ & - 0,003BSW * OP.RFQ - 0,123v \end{aligned} \quad (5)$$

O modelo sugere a inclusão de algumas variáveis significativas para explicar a variável resposta, bem como a interação entre algumas variáveis. Através do ajuste é possível verificar que os poços com alta produção apresentam maiores risco de falha. Com relação à variável RGO observa-se que quanto maior o valor de RGO menor é o tempo livre de falha dos poços. Aparentemente esse resultado parece indicar que a presença de gás afeta consideravelmente o desempenho dos equipamentos de sub-superfície dos poços, levando-os a apresentar maiores risco de falha. A idade do poço mostrou-se relacionada com o tempo de vida dos poços, indicando que os mais jovens apresentam menor probabilidade de sobrevivência em relação aos mais antigos. A Figura 5 mostra as curvas de sobrevivência estimadas através do modelo de regressão Weibull para as interações apontadas como significativas.

Tabela 1. Resultados do ajuste do modelo de regressão Weibull.

Parâmetros	Variáveis explicativas	Estimativa	Erro padrão	P-valor
β_0	(Intercepto)	7,1481	0,5289	< 0,0001
β_1	Produção base	-0,0454	0,0067	< 0,0001
β_2	Valor do BSW	0,0056	0,0023	0,0174
β_3	Método de elevação BM	1,7718	0,5169	0,0006
β_4	Valor do RGO	-0,0028	0,0013	0,0243
β_5	Idade do poço	0,0714	0,0082	< 0,0001
β_6	Unidade Operacional OP-CAM	0,6607	0,3289	0,0446
β_7	Unidade Operacional OP-ET	2,6838	0,7923	0,0007
β_8	Unidade Operacional OP-RFQ	1,1049	0,4749	0,0200
β_9	Profundidade do poço	0,0004	0,0007	0,5230
β_{10}	Profundidade da bomba	0,0018	0,0005	< 0,0001
β_{11}	Método de elevação BM – Unidade Operacional OP-CAM	0,8340	0,3232	0,0098
β_{12}	Método de elevação BM – Unidade Operacional OP-ET	-2,1685	0,7113	0,0023
β_{13}	Método de elevação BM – Unidade Operacional OP-RFQ	-0,0381	0,4694	0,9350
β_{14}	Método de elevação BM – Profundidade do poço	-0,0022	0,0007	0,0030
β_{15}	Valor do BSW – Unidade Operacional OP-CAM	-0,0053	0,0032	0,0971
β_{16}	Valor do BSW – Unidade Operacional OP-ET	-0,0117	0,0042	0,0049
β_{17}	Valor do BSW – Unidade Operacional OP-RFQ	-0,0028	0,0048	0,5640
	Log(scale)	-0,1277	0,0370	0,0006

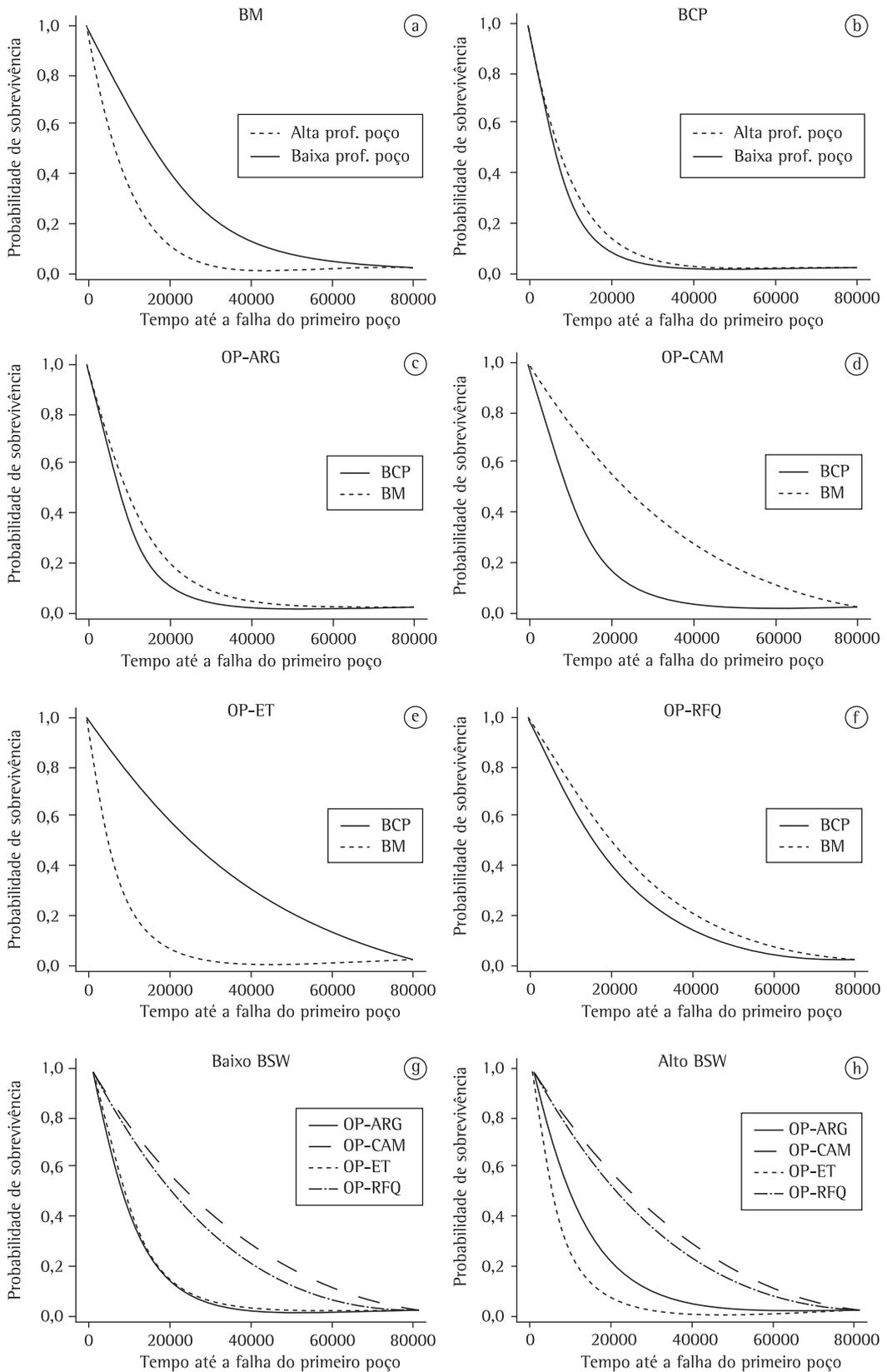


Figura 5. Curva de sobrevivência estimada através do modelo de regressão Weibull.

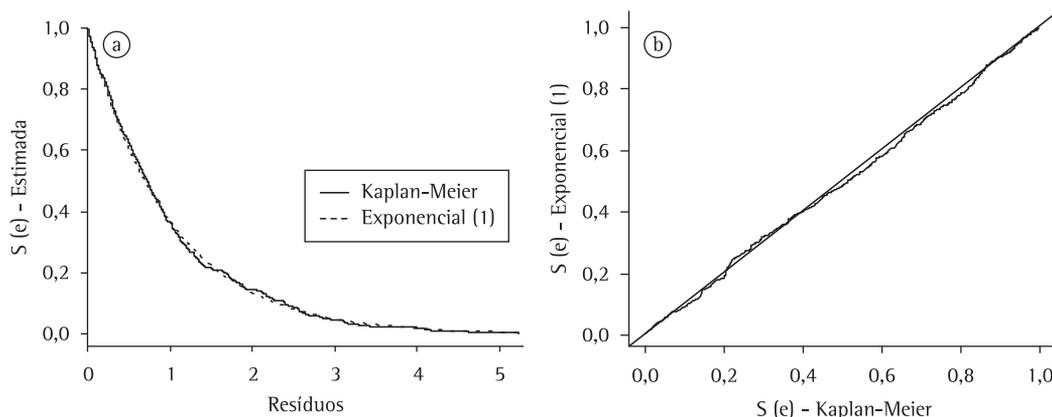


Figura 6. Análise dos resíduos de Cox-Snell do modelo de regressão Weibull ajustado aos dados de petróleo.

Os poços funcionando com método de elevação BM apresentam maior confiabilidade que os poços BCP nas unidades administrativas OP-CAM, na OP-ARG e OP-RFQ. Na unidade operacional OP-ET nota-se que os poços funcionando com método de elevação BCP são mais duráveis que aqueles funcionando com método de elevação BM. E os poços funcionando com esse método e com baixa profundidade apresentam maior tempo de vida em relação aos poços com alta profundidade. Entretanto, a durabilidade dos poços funcionando através do método BCP parece ocorrer o oposto, ou seja, os poços mais profundos apresentam maior tempo de funcionamento.

Nota-se que em média quanto maior o valor da variável BSW maior o tempo de vida dos poços. Quando observado os poços da OP-RFQ e OP-ARG, percebe-se que o tempo de vida cresce à medida que aumenta o volume de água (BSW) e permanece instável na OP-CAM. Esse resultado indica que aparentemente a produção elevada de água não provoca dano aos equipamentos instalados na sub-superfície dos poços. Agora, para os poços da unidade OP-ET, a produção mais elevada de água provoca redução no tempo de vida.

Para confirmar a escolha do modelo usado no ajuste dos dados, foi realizada uma análise gráfica dos resíduos de Cox-Snell (LAWLESS, 1982). Nesta análise são comparadas as curvas de Kaplan-Meier dos resíduos e do modelo exponencial padrão. Na Figura 6 vemos que, pela proximidade das curvas de KM do resíduo e a obtida pelo modelo exponencial padrão, o modelo Weibull apresenta um ótimo ajuste aos dados.

5. Conclusão

Este trabalho realiza 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. O ajuste da distribuição Weibull aos dados dos poços da Bacia Potiguar mostrou-se bastante satisfatório, confirmando o recomendado pela revisão bibliográfica. Através do ajuste, observa-se que os poços produzindo com método de elevação BM são mais duráveis em relação aos poços BCP, para aqueles localizados nas unidades operacionais Canto do Amaro, Fazenda Riacho da Forquilha e Alto do Rodrigues. Os poços com alta produção de óleo tendem a apresentar menor tempo de falha em relação aos menos produtivos. Os poços da unidade Canto do Amaro com baixa e alta produção de BSW mostram maior probabilidade de funcionamento do que os demais poços das outras unidades operacionais. Os poços mais antigos apresentam maior probabilidade de sobrevivência em relação aos poços mais jovens – 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 assim o tempo de vida dele. Altos valores de RGO provocam menor durabilidade dos poços. Em geral, os poços com a bomba localizada em maior profundidade levam mais tempo até apresentar a primeira falha.

Por fim, a partir do estudo efetuado foi possível identificar características intrínsecas dos poços que influenciam o tempo de funcionamento deles, tais como unidade administrativa, profundidade e idade do poço. Além disso, os resultados obtidos com base no modelo estatístico ajustado aos dados amostrais (2000-2006) são uma contribuição para a criação de um sistema estatístico computacional que viabilize o processo decisório na Petrobras, com base em projeções de cenários de riscos de falha em equipamento de sub-superfície.

Referências

- ACCIOLY, R. M. S.; MARTINS, J. A. *Análise do tempo entre as intervenções e duração das intervenções com sonda*. Brasil: PETROBRAS, 1999.
- BARDY, M. B. *Análise de dados de falha de equipamentos de bombeio centrífugo utilizando distribuição de Weibull*. Rio de Janeiro, 2000. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal Fluminense - UFF.
- BOLFARINE, H.; SANDOVAL, M. C. *Introdução à inferência estatística*. Rio de Janeiro: Sociedade Brasileira de Matemática, 2001.
- COLLET, D. *Modelling survival data in medical research*. 1 ed. London: Chapman and Hall, 1994.
- COLOSIMO, E. A.; GIOLO, S. R. *Análise de sobrevivência aplicada*. São Paulo: Editora Edgar Blücher, 2006.
- COX, D. R. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society: Series B*, v. 34, n. 2, p. 187-220, 1972.
- FROTA, H. M. *Desenvolvimento de método para planejamento da manutenção de poços petrolíferos em águas profundas*. Macaé, 2003. Dissertação (Mestrado em Engenharia de Reservatório e de Exploração) - Universidade Estadual do Norte Fluminense - UENF.
- 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.
- KAPLAN, E. L.; MEIER, P. Nonparametric estimation from incomplete observations. *Journal of American Statistical Association*, v. 53, n. 282, p. 457-481, 1958.
- LAWLESS, J. F. *Statistical models and methods for lifetime data*. New York: Wiley, 1982.
- LINQVIST, B.; MOLNES, E.; RAUSAND, M. Analysis of SCSSV performance data. *Reliability Engineering and System Safety*, v. 20, n. 11, p. 3-17, 1988.
- NELSON, W. *Accelerated life testing: statistical models data analysis and test plants*. New York: John Wiley and Sons, 1990.
- PERCY, D. F.; ALKALI, B. M. Generalized proportional intensities models for repairable systems. *Management Mathematics*, v. 17, n. 2, p. 171-185, 2005.
- PERCY, D. F.; ALKALI, B. M. Scheduling preventive maintenance for oil pumps using generalized proportional intensities models. *International Transactions in Operational Research*, v. 14, n. 6, p. 547-563, 2007.

Agradecimentos

À Petrobras pelo fornecimento dos dados deste trabalho e pelo apoio financeiro necessário para o desenvolvimento do Projeto SAFES – Sistema de Análise de Falhas de Equipamentos de Sub-superfície de poço de petróleo da Bacia Potiguar, do qual este trabalho é parte integrante.

Ao engenheiro da Petrobras Dr. Benno Assmann pelo auxílio no entendimento dos dados e a toda equipe envolvida no SAFES.

Weibull-regression models to study failure data in oil pumps

Abstract

This work presents a reliability study conducted among on-shore oil fields in the Potiguar Basin (RN/CE) belonging to PETROBRAS, Brazil. The main study objective was to verify the existence of a relationship between oil pump lifetime and some characteristics such as the elevation method, the amount of water produced in the well (BSW), the gas/oil ratio (RGO), the depth of the production pump, the oil field operational unit, among others. The study was based on a retrospective sample of 450 oil columns from all that were functioning between 2000 and 2006. Statistical hypothesis tests under a Weibull regression model fitted to the first failure data allowed the selection of some covariates in the set considered to explain the failure time in the oil pumps.

Keywords

Reliability. Weibull regression model. On-shore oil fields.