

DESENVOLVIMENTO DE METODOLOGIA BAYESIANA DE ANÁLISE QUANTITATIVA DE RISCO PARA TERMINAIS DE GNL E APLICAÇÃO PARA UM TERMINAL OFFSHORE

Marilia Abílio Ramos

Universidade Federal de Pernambuco – Departamento de Engenharia Química
Av. Prof. Moraes Rego, 1235, CDU, Recife-PE
marilia.aramos@gmail.com

Enrique Andrés López Droguett

Universidade Federal de Pernambuco – Departamento de Engenharia de Produção
Av. Prof. Moraes Rego, 1235, CDU, Recife-PE
ealopez@uol.com.br

Marcelo Ramos Martins

Universidade de São Paulo – Departamento de Engenharia Naval e Oceânica
Av. Prof. Mello Moraes, 2231, CDU, São Paulo-SP
mmartin@usp.br

Celmy Barbosa

Universidade Federal de Pernambuco – Departamento de Engenharia Química
Av. Prof. Moraes Rego, 1235, CDU, Recife-PE
celmy@ufpe.br

RESUMO

O uso do Gás Natural Liquefeito – GNL tem aumentado consideravelmente nas últimas décadas, e com ele a quantidade de terminais de armazenamento e regaseificação. Por causa da natureza inflamável do metano, em caso de acidentes com vazamento de GNL esses terminais podem oferecer perigo às comunidades próximas a eles. Nesse contexto, é essencial a realização da Análise Quantitativa de Risco sobre os terminais. Para cálculo do risco é necessária a quantificação das consequências do acidente e de sua frequência. Entretanto, no caso de indústrias relativamente recentes, como GNL, os dados de frequência de acidentes e falhas de equipamentos são esparsos, não sendo estatisticamente confiáveis. A análise Bayesiana em dois estágios permite lidar com este tipo de situação, pois permite o uso de dados específicos da planta estudada juntamente com informações obtidas em outras indústrias, com mais tempo de existência e portanto dados de taxas de falha de equipamentos mais confiáveis.

Análise Bayesiana em Dois Estágios, Análise Quantitativa de Risco, Terminal de GNL

ABSTRACT

The use of Liquefied Natural Gas – LNG has been increasing at a considerable rate, and so the quantity of its storage and regasification terminals. Because of the flammable nature of methane, in cases of accident, the terminals can endanger communities located around them. In this context, it is essential to do Quantitative Risk Analysis for the terminals. In order to perform it, it is necessary to quantify the consequences of the accident and its frequency. In case of relatively recent industries, such as LNG industry, however, data regarding the frequency of accidents and equipment failures are sparse, thus not statistically reliable. The Two Stages Bayesian Analysis has proven to be an outstanding tool to deal with this kind of situation, because it allows the use of data specific to the studied plant in conjunction with information from other industries, usually older and with more reliable data regarding equipment failure rate.

Two Stages Bayesian Analysis, Quantitative Risk Analysis, LNG Terminal

1. Introdução

O gás natural é uma fonte de energia fóssil composta quase totalmente por metano, podendo ainda conter pequenas frações de etano, propano, gás carbônico e outros. Apesar de suas vantagens em relação a outros combustíveis fósseis, especialmente a menor quantidade de gases poluentes emitidos durante sua queima, o gás natural foi visto durante muito tempo como subproduto indesejado da exploração de petróleo.

Nas últimas décadas, entretanto, este cenário tem se alterado. Em 2008, o gás natural tornou-se a segunda maior fonte de energia elétrica no mundo, superado apenas pelo carvão (IEA, 2008). No Brasil, particularmente, a participação do gás natural na matriz energética é de 10,3% (MME, 2011), comprovando a atuação mais forte que vem exercendo também em âmbito nacional.

Há, contudo, entraves para o crescimento do uso do gás natural. Um deles é a dificuldade existente para seu transporte e armazenamento, especialmente devido ao seu estado físico gasoso. O Gás Natural Liquefeito - GNL é uma tecnologia que permite superar essas dificuldades, por causa tanto de seu estado físico líquido como também por reduzir o volume do gás natural em cerca de 600 vezes durante o processo de liquefação. Dadas estas vantagens, o GNL é uma das formas mais utilizadas para transportar e armazenar o gás natural atualmente.

A cadeia de valor dessa tecnologia consiste de quatro etapas interdependentes: exploração e produção do gás natural, sua liquefação, o transporte através de navios ou caminhões-tanque, sua estocagem e a regaseificação para envio ao consumidor final (CEE, 2007). É importante ressaltar que os terminais de estocagem e regaseificação do GNL podem ser *onshore* ou *offshore*, e estes últimos, por sua vez, podem ser do tipo *Gravity Based System*, constituído de uma estrutura de concreto com base no fundo do mar, e *Floating Storage and Regasification Unit* (FSRU), navios com tanques de GNL e unidade de regaseificação *onboard*, sendo estes últimos uma tendência mundial.

O Gás Natural Liquefeito tem sido visto com crescente interesse pelo Brasil. Em 2009 já existiam dois terminais do tipo FSRU em operação, um em Pecém-CE e o outro na Baía de Guanabara-RJ, e mais um deve entrar em operação nos próximos anos, na Baía de Todos os Santos – BA. Os terminais *offshore* apresentam vantagens sobre os terminais *onshore*, tais como flexibilidade, menor custo e maior rapidez no retorno de capital investido (BLACKWELL; SKAAR, 2009). Outro benefício apresentado pelos terminais *offshore* é a sua distância das áreas urbanas, que pode reduzir os riscos às populações em caso de acidente com vazamento do produto (BLACKWELL; SKAAR, 2009; WIJNGAARDEN; OOMEN; HOORN, 2004).

A importância que o GNL tem conquistado no cenário energético nos últimos anos, especialmente no Brasil, e o consequente aumento do número de terminais de armazenamento e regaseificação, geram preocupações quanto aos riscos a que estão expostas pessoas próximas às instalações. Apesar de a indústria de GNL ter registrado baixo número de acidentes nos últimos 40 anos, os riscos relacionados aos terminais podem estar crescendo com o aumento de sua quantidade (YUN, 2007). Neste contexto, é essencial analisar e quantificar estes riscos, de forma a manter o baixo registro de acidentes da indústria e permitir, assim, que se continue aproveitando do GNL para transporte e armazenamento do gás natural.

Os principais perigos colocados por um acidente com vazamento de GNL em um terminal são relacionados a incêndio (WOODWARD; PITBLADO, 2010). Na existência de uma fonte de ignição no momento do vazamento, a poça formada é inflamada resultando em um incêndio. Não ocorrendo ignição imediata, há a formação de uma nuvem de vapor que pode sofrer ignição se dentro dos limites de inflamabilidade do gás natural. Por causa de sua baixa temperatura, partículas de água presentes na atmosfera condensam dentro da nuvem, deixando-a densa e com baixa altitude. O fogo do incêndio em nuvem pode, portanto, entrar em contato com a poça em evaporação e causar um incêndio em poça retardado (LUKETA-HANLIN, 2005).

As radiações emitidas pelos incêndios em poça e em nuvem são relacionadas à probabilidades de fatalidade – vulnerabilidade – das pessoas expostas (CETESB, 2003). Através de uma análise de consequências, é possível prever a distância atingida por certo nível de radiação e consequente vulnerabilidade das pessoas atingidas. Combinando-se estes resultados

com a frequência com que ocorre o acidente, tem-se o risco ao qual as pessoas estão expostas (AICHE, 2000).

As análises de consequência e vulnerabilidade e frequência dos acidentes constituem a Análise Quantitativa de Risco (AQR), que permite prever o risco ao qual estão expostas as pessoas na vizinhança do empreendimento (CAMACHO, 2004). A aplicação da AQR aos terminais de estocagem e regaseificação de GNL é de grande importância, em especial no contexto do Brasil, cujo interesse pelos terminais tem se provado grande.

O cálculo da frequência dos acidentes, entretanto, é uma etapa cuja metodologia ainda não está bem estabelecida para as indústrias de GNL. De uma forma geral, a frequência pode ser calculada através de bancos de dados “genéricos” de falhas de equipamentos, referentes a indústrias em geral, ou dados “específicos”, provenientes da planta estudada (KAPLAN, 1983). Entretanto, no caso de um terminal de GNL, o uso exclusivo de um destes tipos de dados no cálculo da frequência de um evento não fornece resultados realistas, pois os dados “genéricos” são geralmente obtidos em condições diferentes das da indústria de GNL e os dados “específicos” são escassos devido ao pouco tempo de existência da indústria de GNL (YUN, 2007, 2008).

Nesse contexto, é necessário o desenvolvimento de uma metodologia para o cálculo de frequências dos acidentes na indústria de GNL, de forma que seja possível se beneficiar da estabilidade que têm os dados de indústrias similares e a adequabilidade dos dados específicos. A metodologia proposta nesse trabalho é a Análise Bayesiana em Dois Estágios. O presente artigo tem, assim, o objetivo de apresentar esta metodologia e desenvolver uma aplicação da mesma em um terminal *offshore* de GNL.

2. Análise Bayesiana

Tendo por base o teorema de Bayes, a análise Bayesiana consiste na utilização de informações disponíveis, as chamadas evidências, em uma função chamada verossimilhança, para atualizar um conhecimento inicial sobre a distribuição de um parâmetro de interesse – distribuição a priori. Esta atualização resulta em uma distribuição a posteriori do parâmetro estudado (KELLY; SMITH, 2011).

Assim, o teorema de Bayes provê um método matemático para atualização de um conhecimento sobre determinado evento a partir da introdução de evidências sobre o mesmo. De acordo com o teorema de Bayes:

$$p(\theta|y) = \frac{p(\theta) \times p(y|\theta)}{p(y)} \quad (\text{Eq. 1})$$

sendo:

$p(\theta|y)$ é a distribuição a posteriori, que expressa o conhecimento sobre o parâmetro θ após a obtenção das evidências y ;

$p(\theta)$ é a distribuição a priori do parâmetro θ , que expressa o conhecimento sobre θ antes da obtenção das evidências;

$p(y|\theta)$ é a função de verossimilhança ou distribuição de probabilidade dos dados y condicionais nos parâmetros θ , que fornece a plausibilidade de cada um dos possíveis valores do parâmetro θ ;

$p(y)$ é a distribuição marginal, e age como uma constante de normalização.

O Teorema de Bayes pode ser escrito para a variáveis contínuas, conforme Equação 2 :

$$\pi(\theta|y) = \frac{f(y|\theta) \times \pi_0(\theta)}{\int f(y|\theta) \times \pi_0(\theta) d\theta} \quad (\text{Eq. 2})$$

A escolha da distribuição a priori é uma etapa subjetiva da análise. Ela pode ser informativa ou não informativa. As distribuições informativas contêm informações substanciais sobre os possíveis valores de θ enquanto as não informativas são utilizadas quando pretende-se deixar que as evidências dominem a distribuição a posteriori, ou seja, contêm pouca informação sobre o parâmetro de interesse (GREGORY, 2005).

A abordagem Bayesiana apresenta diversas vantagens sobre a abordagem clássica, dentre as quais podem-se enumerar as seguintes:

- Permite que se combinem dados a uma informação anterior a esses através de um modelo matemático. Quando novas informações tornam-se disponíveis, a distribuição a posteriori pode ser utilizada como a priori em uma nova análise;
- Permite o uso de diversos tipos de dados como evidência;
- Redução de Custos: estimativas e *feedback* baseados apenas em dados de teste são onerosos e podem trazer apenas benefícios marginais. Métodos Bayesianos usam a totalidade do conhecimento disponível e não apenas dados estatísticos. Acumular a mesma quantidade de informação através de testes é frequentemente inviável principalmente para produtos altamente confiáveis.

Dentre as desvantagens da análise Bayesiana, podem ser citados o maior esforço para sua compreensão e aplicação, além do caráter subjetivo empregado na identificação e justificativa da distribuição a priori.

A respeito dos tipos de dados a serem tratados em uma análise Bayesiana, ela pode pertencer a duas classificações distintas: homogênea ou não homogênea. A primeira é utilizada quando a população é homogênea, ou seja, a população é idêntica em relação a algumas características de interesse. Se, por exemplo, todos os membros de uma população possuem a mesma taxa de falha, ela é dita homogênea em relação à taxa de falha. No caso de subpopulações homogêneas, os dados utilizados como verossimilhança podem ser agrupados e utilizados em uma única aplicação do teorema de Bayes para se obter o conhecimento a posteriori representativo de toda a informação acumulada.

Já no caso de uma análise Bayesiana não homogênea, lida-se com subpopulações que exibem características distintas entre si, devido a fatores como *design*, diferenças operacionais ou condições ambientais, por exemplo. Neste caso, não é realista assumir que todos os membros da população, composta das diferentes subpopulações, possuem os mesmos parâmetros de confiabilidade, como taxa de falha. A característica de interesse, no caso de uma população não-homogênea, é inerentemente diferente de uma subpopulação para outra. Isto é chamado variabilidade populacional da característica em questão. A variabilidade populacional faz parte de uma população não homogênea, e não pode ser eliminada com maior quantidade de informações. Para avaliar a distribuição a posteriori através de uma análise Bayesiana não homogênea é necessário encontrar uma distribuição desta variabilidade populacional para o parâmetro, digamos uma taxa de falha λ : $\phi(\lambda)$.

É possível assumir que a distribuição da variabilidade populacional $\phi(\lambda)$ é membro de família paramétrica, como beta, gama, lognormal. Sendo $\underline{\theta} = \{\theta_1, \theta_2, \dots, \theta_m\}$ o conjunto dos parâmetros de $\phi(\lambda)$, ou seja, $\phi(\lambda) = \phi(\lambda|\underline{\theta})$, a distribuição da incerteza, ou estado de conhecimento, sobre $\phi(\lambda)$ é igual à sobre os valores de $\underline{\theta}$. Para cada valor de $\underline{\theta}$ existe um único $\phi(\lambda|\underline{\theta})$. Assim, estimando-se $\underline{\theta}$ estima-se também $\phi(\lambda)$.

Dada a evidência disponível, E, a uma distribuição a priori de $\underline{\theta}$, o Teorema de Bayes pode ser utilizado para encontrar a distribuição de probabilidade do estado de conhecimento de $\underline{\theta}$, de acordo com a Equação 3:

$$\pi(\underline{\theta}|E) = \frac{L(E|\underline{\theta})\pi_0(\underline{\theta})}{\int_{\theta_1} \int_{\theta_2} \dots \int_{\theta_m} L(E|\underline{\theta})\pi_0(\underline{\theta})d\underline{\theta}} \quad (\text{Eq. 3})$$

sendo:

$\pi_0(\underline{\theta})$ é a distribuição a priori de $\underline{\theta}$;

$\pi(\underline{\theta}|E)$ é a distribuição a posteriori de $\underline{\theta}$ dada a evidência E (distribuição conjunta de probabilidade m-dimensional sobre os valores de $\theta_1, \theta_2, \dots, \theta_m$);

$L(E|\underline{\theta})$ é a verossimilhança de E dada $\underline{\theta}$.

A Figura 1 exemplifica a distribuição conjunta de probabilidade dos parâmetros θ_1, θ_2 :

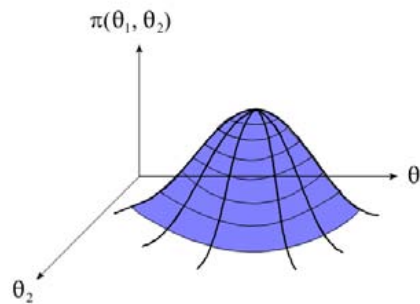


Figura 1: Exemplo de distribuição conjunta de probabilidade de $\underline{\theta} = \{\theta_1, \theta_2\}$

A distribuição esperada de probabilidade da variabilidade populacional na análise Bayesiana não homogênea é dada, finalmente, por:

$$\bar{\phi}(\lambda) = \int_{\theta_1} \dots \int_{\theta_m} \phi(\lambda|\theta_1 \dots \theta_m)\pi(\theta_1 \dots \theta_m|E)d\theta_1 \dots \theta_m \quad (\text{Eq. 4})$$

3. Metodologia Bayesiana de cálculo de frequências

As taxas e frequências de falha de equipamentos em uma Análise Quantitativa de Risco são, em geral, estimadas a partir de dados empíricos como número de falhas de um equipamento e o tempo total de exposição, ou a partir de opiniões de especialistas ou estimativas disponíveis em bancos de dados genéricos de confiabilidade. Existem dois tipos de fontes destes dados de falha: dados coletados na própria planta que está sendo estudada, chamados de dados específicos, e dados reportados por indústrias em geral, chamados de dados genéricos.

À primeira vista, parece mais correta a utilização dos dados específicos na Análise de Risco. Entretanto, em diversas situações, estes dados não são estatisticamente confiáveis ou simplesmente não existem. Os dados genéricos, por outro lado, podem não caracterizar muito bem o equipamento em questão, visto que a falha de um equipamento é bastante dependente de suas condições de operação, manutenção, fornecedores, etc. (RAMOS; 2011). Assim, os dois tipos de fontes de dados possuem limitações em seu uso, especialmente no caso de indústrias com pouco tempo de funcionamento ou no estudo de eventos muito raros, que podem, até mesmo, ser ainda hipotéticos no momento da realização do estudo.

Através do uso da análise Bayesiana é possível utilizar os dois tipos de dados no cálculo das frequências, considerando assim a experiência histórica de outras indústrias porém parcialmente relevantes à planta em análise, e experiências específicas da planta em questão. Além da possibilidade de utilizar dados genéricos e específicos, podem ser utilizadas também diversas fontes de dados, como opiniões de especialistas e dados de falha no tempo (empíricos) (SIU; KELLY,1998; SHAFAGHI,2006).

Por estas razões, a análise Bayesiana é bastante aplicável em indústrias de Gás Natural Liquefeito. O tempo de operação das indústrias de GNL não é suficiente para garantir relevância estatística aos dados de falha, ao contrário de indústrias como refinarias e outras petroquímicas. Assim sendo, o cálculo de frequências dos acidentes em um terminal de GNL pode ser feito com uso da análise Bayesiana, de forma a acrescentar à base de dados de falhas de equipamentos da indústria de GNL, os dados genéricos de outras indústrias, em funcionamento há mais tempo, assim como opiniões de especialistas e estimativas de bancos de dados de confiabilidade.

Nos métodos Bayesianos, os dados genéricos são geralmente modelados como a distribuição a priori e os dados específicos como as funções de verossimilhança (YAN; HAIMES,2010). Esta aplicação convencional do teorema de Bayes foi utilizada por Yun (2007) para determinação das frequências de acidentes em um terminal de GNL, utilizando dados genéricos de bancos de dados como OREDA (*Offshore Reliability Data*) e EIREDA (*European*

Industry Reliability Data Bank) como distribuição a priori e atualizando-os com dados de plantas de GNL.

Contudo, esta metodologia tem um problema: apesar de os dados serem coletados a partir de uma variedade de fontes, como diferentes plantas com diferentes condições de operação, eles são tratados como provenientes de uma mesma fonte, ou seja, supõe-se que sejam constituintes de uma população homogênea. As incertezas provenientes da variabilidade de fonte para fonte são ignoradas (como as diferenças entre as condições operacionais, políticas de manutenção, diferenças de fabricantes, etc.), o que faz com que a distribuição a posteriori resultante tenda a ser muito estreita, ou seja, ocorre uma subestimação do nível de incerteza (SIU, KELLY,1998). Por essas razões, a metodologia aplicada por Yun (2007,2008) para a determinação de frequências dos acidentes em um terminal de GNL é inapropriada quando aplicada para contextos com escassez de dados e o uso de fontes alternativas de informação.

A Análise Bayesiana em Dois Estágios, primeiramente introduzida no contexto de análise de risco por Kaplan (1983), considera esta não-homogeneidade dos dados de falha na estimativa dos parâmetros dos modelos probabilísticos de uma determinada grandeza de interesse (como a frequência de ocorrência de um evento acidental), cuja distribuição a posteriori resultante é então utilizada como a distribuição a priori no segundo estágio da análise, conforme Figura 2.

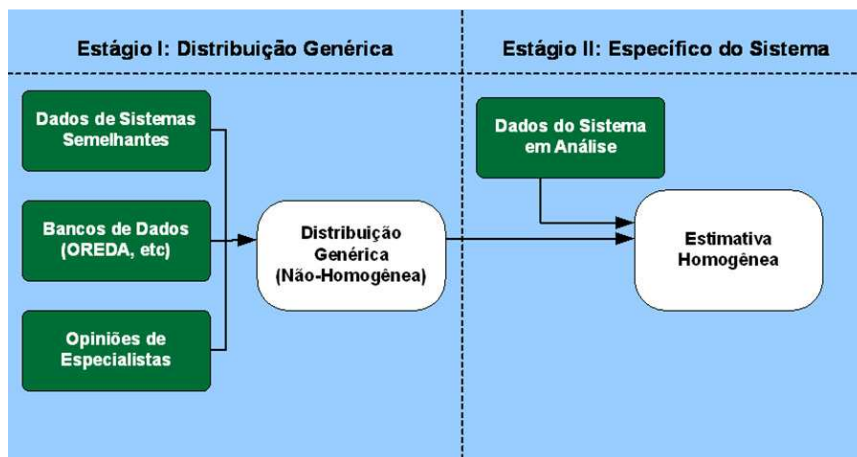


Figura 2: Análise Bayesiana em Dois Estágios

Devido ao curto tempo de existência dos terminais de GNL, à conseqüente falta de dados estatisticamente relevantes sobre estes terminais e à não-homogeneidade inerente aos dados disponíveis, a Análise Bayesiana em Dois Estágios é uma forma apropriada para calcular a frequência de ocorrência de um evento em um terminal de GNL, e pode ser incluída na metodologia de AQR para plantas de GNL.

Durante o primeiro estágio da análise, o objetivo é desenvolver a distribuição de probabilidade caracterizando a incerteza sobre os parâmetros da variabilidade populacional $\pi(\underline{\theta}|E)$ (seguindo a premissa de que ela faz parte de uma família paramétrica de distribuições), a partir da qual poderá ser gerada uma distribuição média (esperada) da variabilidade populacional que poderá então ser utilizada como a distribuição a priori no segundo estágio da análise.

Considerando que o conhecimento inicial sobre os parâmetros $\underline{\theta}$ é representado pela distribuição de probabilidade $\pi_0(\underline{\theta})$, e dada a evidência disponível E, a distribuição de probabilidade a posteriori de $\underline{\theta}$ é dada pelo teorema de Bayes conforme mostrado na Equação 5.

$$\pi(\underline{\theta}|E) = \frac{L(E|\underline{\theta})\pi_0(\underline{\theta})}{\int_{\theta_1} \int_{\theta_2} \dots \int_{\theta_m} L(E|\underline{\theta})\pi_0(\underline{\theta})d\underline{\theta}} \quad (\text{Eq. 5})$$

O conjunto de evidências E do parâmetro de interesse X pode ser composto por quatro tipos de evidências:

E_0 : Estado de conhecimento inicial do analista (esta evidência é utilizada na construção da distribuição a priori, e pode indicar mesmo que o conhecimento inicial do analista é muito pouco ou nenhum, gerando a priori não-informativa);

E_1 : Dados empíricos (quantidade de falhas por tempo ou por demanda) de sistemas similares operando em condições similares;

E_2 : Estimativas ou distribuições sobre X a partir de bancos de dados de confiabilidade ou opiniões de especialistas;

E_3 : Dados empíricos específicos do sistema analisado.

A distribuição do parâmetro de interesse, depois de obtidas as evidências E_0, E_1, E_2 e E_3 , pode ser escrita então como:

$$\pi(x|E_0, E_1, E_2, E_3) = \frac{P(E_3|x, E_0, E_1, E_2) \int_{\underline{\theta}} \varphi(x|\underline{\theta}) \pi(\underline{\theta}|E_0, E_1, E_2) d\underline{\theta}}{\int_x P(E_3|x, E_0, E_1, E_2) \int_{\underline{\theta}} \varphi(x|\underline{\theta}) \pi(\underline{\theta}|E_0, E_1, E_2) d\underline{\theta} dx} \quad (\text{Eq. 6})$$

em que a distribuição a priori, $\int_{\underline{\theta}} \varphi(x|\underline{\theta}) \pi(\underline{\theta}|E_0, E_1, E_2) d\underline{\theta}$, é a densidade da variabilidade populacional esperada, obtida como distribuição a posteriori no primeiro estágio da Análise Bayesiana em Dois Estágios.

A Equação 6 é utilizada apenas no segundo estágio, quando tem-se dados específicos da planta analisada.

Explicitando as evidências consideradas, a distribuição dos parâmetros $\pi(\underline{\theta}|E)$ mostrada na Equação 5 pode ser escrita como:

$$\pi(\underline{\theta}|E_0, E_1, E_2) = \frac{P(E_1, E_2|\underline{\theta}, E_0) \pi_0(\underline{\theta}|E_0)}{\int_{\underline{\theta}} P(E_1, E_2|\underline{\theta}, E_0) \pi_0(\underline{\theta}|E_0) d\underline{\theta}} \quad (\text{Eq. 7})$$

Assumindo que as evidências E_1 e E_2 são independentes entre si, e considerando que normalmente existe mais de uma fonte para cada tipo de evidência, a verossimilhança da Equação 7 pode ser escrita como

$$P(E_1, E_2|\underline{\theta}, E_0) = \prod_{i=1}^n P(E_{1i}|\underline{\theta}, E_0) P(E_{2i}|\underline{\theta}, E_0) \quad (\text{Eq. 8})$$

onde $P(E_{1i}|\underline{\theta}, E_0)$ e $P(E_{2i}|\underline{\theta}, E_0)$ são as probabilidades de observar as evidências E_{1i} e E_{2i} , respectivamente, para i -ésimo de n sistemas, assumindo que o parâmetro da curva de variabilidade populacional é $\underline{\theta}$.

A medida exata do parâmetro de interesse em cada sistema, x_i , é desconhecida, sabendo-se apenas que ele é um dos possíveis valores de X . Além disso, X é distribuído de acordo com $\varphi(x|\underline{\theta})$, com $\underline{\theta}$ também desconhecido. Assim, calcula-se a probabilidade de observar as informações E_1 e E_2 considerando que o parâmetro de interesse pode assumir todos os valores possíveis, ou seja, integrando $P(E_1, E_2|\underline{\theta}, E_0)$ sobre toda a distribuição de X :

$$P(E_{1i}, E_{2i}|\underline{\theta}, E_0) = \int_x P(E_{1i}|\underline{\theta}, E_0) P(E_{2i}|\underline{\theta}, E_0) \varphi(x|\underline{\theta}) dx \quad (\text{Eq. 9})$$

A Equação 9 pode, então, ser substituída na Equação 8 para obter a função de verossimilhança a partir de todas as evidências.

No caso de a evidência disponível ser apenas do tipo E_1 , a verossimilhança será $\int_x P(E_{1i}|\underline{\theta}, E_0) \varphi(x|\underline{\theta}) dx$, e, de forma similar, possuindo apenas evidência do tipo E_2 a verossimilhança será $\int_x P(E_{2i}|\underline{\theta}, E_0) \varphi(x|\underline{\theta}) dx$.

Os modelos de probabilidade para verossimilhança baseada em dados empíricos e em opiniões de especialista podem ser vistos em Droguett e Mosleh (2005), e para a verossimilhança composta dos dois tipos de informações (verossimilhança mista) em Droguett, Groen e Mosleh, (2003).

4. Aplicação da metodologia em uma terminal offshore de GNL

A metodologia Bayesiana em Dois Estágios para cálculo das frequências de acidentes na indústria de GNL foi aplicada para um terminal hipotético *offshore* do tipo FSRU. O cálculo de frequências foi precedido da realização de uma Análise Preliminar de Perigo, em que são identificados os possíveis acidentes com vazamento de produto, suas possíveis causas e consequências, assim como classificação quanto à severidade. Foram identificados, no total, 44 cenários, entre vazamentos em linha e ruptura de equipamentos. Destes, seis foram considerados como mais críticos e, assim, representativos dos demais. Os cenários representativos, chamados Eventos Iniciadores (EIs) ou Hipóteses Acidentais, podem ser vistos na Tabela 1. Como exemplo da aplicação da Metodologia Bayesiana de para o cálculo de frequência na AQR será explorado nesse artigo o Evento Iniciador 02.

Tabela 1: Eventos Iniciadores identificados na Análise Preliminar de Risco

EI	Perigo	Trecho
01	Grande liberação de GNL	Braço de descarregamento de GNL
02	Grande liberação de GNL	Transferência de líquido da costa até os tanques de GNL
03	Grande liberação de gás natural	GN da estação de medição até o limite de bateria
04	Grande liberação de gás natural	Transferência de vapor dos tanques de GNL até o compressor
05	Grande liberação de GNL	Tubulação no cais entre as válvulas de emergência
06	Grande liberação de GNL	Tanque de armazenamento de GNL

A análise de frequências foi realizada até o primeiro estágio da Análise Bayesiana em Dois Estágios, já que o terminal considerado é hipotético e portanto não possui dados “específicos” de falhas de equipamentos. Esta situação é típica de AQR de plantas em fase de concepção onde não há experiência operacional.

O Evento Iniciador 02 refere-se à ruptura da linha de transferência de Gás Natural Liquefeito do navio provedor aos tanques de GNL no terminal. Os dados utilizados neste evento, obtidos em Bancos de Dados e referentes a ruptura de linha, encontram-se na Tabela 2.

Tabela 2: Evidências utilizadas na análise do Evento Iniciador 2

Banco de Dados	Tempo de Operação (ano)	Falhas	Frequência (ano ⁻¹)
LNG (JOHNSON, WELKER; 2001)	$1,329 \times 10^5$	2	-
Hawksley (1984)	-	-	$1,14 \times 10^{-5}$
Purple Book (TNO, 2005)	-	-	$3,80 \times 10^{-5}$
COVO (COVO, 1982)	-	-	$3,34 \times 10^{-5}$
CCPS (CCPS, 1989)	-	-	$5,54 \times 10^{-5}$

Com o intuito de fazer com que a distribuição a posteriori reflita principalmente as evidências coletadas, escolheu-se utilizar uma distribuição a priori não informativa. A distribuição a priori escolhida foi a distribuição lognormal com fator de erro α igual a 100, de forma que ela terá pouco ou nenhum impacto sobre a distribuição a posteriori, ou seja, tem-se uma distribuição com elevada variância e, portanto, difusa. Para localizar a distribuição a priori em torno da média das estimativas obtidas pelos especialistas, a mediana da distribuição a priori

foi estabelecida como sendo a média das frequências dos dados das evidências. A mediana μ da distribuição a priori utilizada no EI02 foi de $1,52 \times 10^{-4}$ ano⁻¹.

A aplicação da análise Bayesiana em dois estágios foi realizada através do software R-DAT Plus v1.5, software desenvolvido especialmente para a realização de análises Bayesianas no contexto da análise de risco.

O primeiro passo da análise é a obtenção da distribuição conjunta de probabilidade dos parâmetros mediana e fator de erro da variabilidade populacional $\pi(\mu, \alpha|E)$, ilustrada na Figura 3.

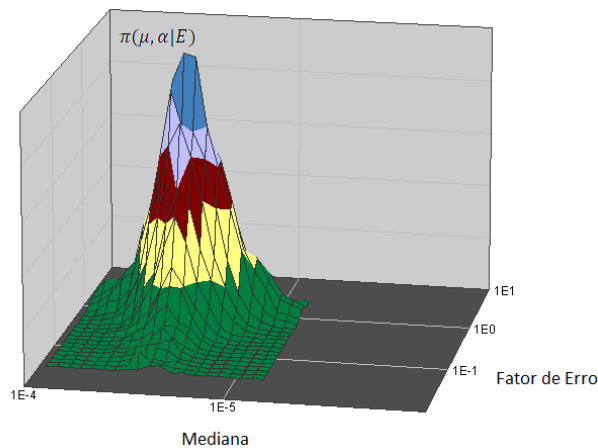


Figura 3: Distribuição a posteriori dos parâmetros da variabilidade populacional - EI02

Com a distribuição conjunta dos parâmetros da variabilidade populacional $\pi(\mu, EF|E)$ e a distribuição lognormal da variabilidade $\phi(p|\mu, \sigma)$, a distribuição esperada da variabilidade populacional da frequência do EI01 pode ser estimada através da equação:

$$\bar{\phi}(\lambda) = \int_{\theta_1} \dots \int_{\theta_m} \phi(\lambda|\theta_1 \dots \theta_m) \pi(\theta_1 \dots \theta_m|E) d\theta_1 \dots \theta_m$$

na forma:

$$\bar{\phi}(\lambda) = \int_{\mu} \int_{\alpha} \phi(\lambda|\mu, \alpha) \cdot \pi(\mu, \alpha) d\mu d\alpha$$

A distribuição acumulada e a função densidade de probabilidade da variabilidade populacional esperada podem ser vistas nas Figuras 4 e 5, respectivamente.

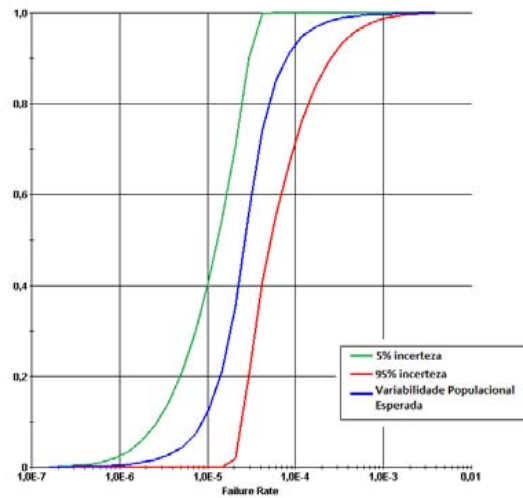


Figura 4: Distribuição acumulada esperada da variabilidade populacional da frequência do Evento Iniciador EI02

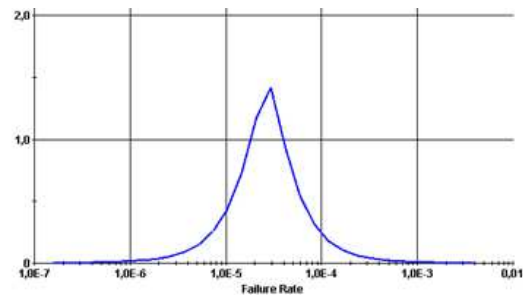


Figura 5: Função densidade de probabilidade da variabilidade populacional esperada das frequências do Evento Iniciador EI02

Na Figura 4 é possível ver que o intervalo de probabilidade – distância entre as curvas de 5% e 95% de confiança – não é significativamente extenso, resultado da quantidade razoável de evidências utilizadas para a verossimilhança.

A Tabela 3 apresenta os valores para as curvas da média e de 5%, 50% e 95% de probabilidade para a variabilidade populacional da frequência do evento.

Tem-se, assim, que a frequência do Evento EI01, determinada pela média da distribuição esperada da variabilidade populacional, é igual a $5,38 \times 10^{-5} \text{ano}^{-1}$. Essa seria a distribuição utilizada como a priori no segundo estágio da análise Bayesiana em dois estágios.

Tabela 3: Média e curvas percentuais da distribuição de variabilidade da frequência do EI02

Valor	Média	Curva de 5%	Curva de 50%	Curva de 95%
Média	$5,38 \times 10^{-5}$			
Variância	$4,60 \times 10^{-2}$			
1st	$1,80 \times 10^{-6}$	$2,01 \times 10^{-5}$	$5,49 \times 10^{-6}$	$6,09 \times 10^{-7}$
5th	$5,59 \times 10^{-6}$	$2,31 \times 10^{-5}$	$8,84 \times 10^{-6}$	$1,63 \times 10^{-6}$
10th	$8,93 \times 10^{-6}$	$2,54 \times 10^{-5}$	$1,14 \times 10^{-5}$	$2,70 \times 10^{-6}$
50th	$2,69 \times 10^{-5}$	$5,18 \times 10^{-5}$	$2,69 \times 10^{-5}$	$1,32 \times 10^{-5}$
90th	$7,75 \times 10^{-5}$	$2,54 \times 10^{-4}$	$6,21 \times 10^{-5}$	$2,96 \times 10^{-5}$
95th	$1,23 \times 10^{-4}$	$4,19 \times 10^{-4}$	$7,96 \times 10^{-5}$	$3,25 \times 10^{-5}$
99th	$3,78 \times 10^{-4}$	$1,15 \times 10^{-3}$	$1,28 \times 10^{-4}$	$3,77 \times 10^{-5}$

Observa-se a diferença entre a frequência encontrada, que corresponde à média da variabilidade populacional da frequência de falha entre os bancos de dados utilizados, e os próprios dados utilizados.

A análise de frequências é uma etapa fundamental da Análise Quantitativa de Risco e é essencial que as frequências sejam bem calculadas. Assim, especialmente em indústrias que possuem dados esparsos de falhas de equipamentos, como a do Gás Natural Liquefeito, a Análise Bayesiana em Dois Estágios deve ser considerada para cálculo final das frequências, afim de que o risco calculado seja o mais realista possível.

5. Conclusões e trabalhos futuros

Apesar de a indústria de Gás Natural Liquefeito não apresentar uma grande quantidade de acidentes envolvendo fatalidades em sua história, é necessário que seus riscos sejam considerados e estudados, visto o perigo em que colocam suas instalações em caso de vazamento de GNL. Desta forma, uma análise dos riscos gerados pelas plantas de GNL deve ser especialmente aplicada aos terminais de armazenamento e regaseificação, devido ao crescimento rápido da quantidade dessas instalações no mundo - especialmente os terminais *offshore*, inclusive no Brasil.

Enquanto diversos estudos tratam das consequências de um vazamento de GNL, com revisão de resultados experimentais e validação de modelos, a metodologia a ser empregada na análise de frequências de um acidente em um terminal não é completamente consolidada. Afinal, devido ao pouco tempo de existência da indústria de GNL, os dados de falhas específicos dessa indústria não são estatisticamente confiáveis. O uso de dados de indústrias similares também não levaria a um resultado correto, por causa das diferenças que existem entre as condições de operação entre as indústrias.

O uso da análise Bayesiana supera essas dificuldades, pois permite a utilização de dados específicos e genéricos, assim como dados empíricos e opiniões de especialistas. O teorema de Bayes, entretanto, não pode ser utilizado de forma direta, utilizando-se os dados específicos da planta como evidências para atualização de um conhecimento inicial, formado pelos dados genéricos, pois os dados são obtidos em diferentes condições de operação, o que torna a população não-homogênea em relação à taxa de falha. O cálculo das frequências de acidentes em terminais, desse modo, subestima a incerteza em relação aos dados, levando a uma distribuição a posteriori dos valores das frequências mais estreita do que realmente é.

Assim, a Análise Bayesiana em Dois Estágios é uma forma adequada para calcular frequências de acidentes em um terminal de GNL, por considerar a não-homogeneidade dos dados. A metodologia foi aplicada para um terminal *offshore* de GNL e foi possível calcular a variabilidade populacional da frequência de falha para um acidente considerado - ruptura da linha de transferência de Gás Natural Liquefeito do navio provedor aos tanques de GNL no terminal. Este resultado pode ser utilizado posteriormente no segundo estágio da análise Bayesiana, quando forem disponíveis dados específicos do terminal. Ainda, ressalta-se que o resultado obtido pode ser combinado à uma análise de consequências, de forma a calcular os riscos social e individual.

6. Agradecimentos

Agradecimentos à ANP que, através do Programa de Recursos Humanos PRH-28, permitiu a realização deste trabalho.

Referências

American Institute of Chemical Engineers – AICHE. *Guidelines for Chemical Process Quantitative Risk Analysis*. 2 ed. New York, 2000.

BLAKWELL, B.; SKAAR, H. *Delivering The World's First FSRUs*. In: 24th WORLD GAS CONFERENCE. Buenos Aires, 2009

CAMACHO, E. *Uma Proposta de Metodologia para Análise Quantitativa de Riscos Ambientais*. 140 folhas. Tese – Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2004.

CCPS. *Guidelines for Process Equipment Reliability Data*, 1989

- CEE - Center for Energy Economics.** *An Overview on Liquefied Natural Gas (LNG), its Properties, Organization of the LNG Industry and Safety Considerations.* Austin, 2007
- CETESB - Companhia Ambiental do Estado de São Paulo.** *Manual de Orientação para Elaboração de Estudos de Análise de Riscos,* P4.261. 2003
- COVO. Rijnmond Public Authority,** *A Risk Analysis of Six Potentially Hazardous Industrial Objects in the Rijnmond Area- A Pilot Study,* COVO, D. Reidel Publishing Co., Dordrecht, 1982
- DROGUETT, E.; MOSLEH, F.** *Bayesian Assessment of the Variability of Reliability Measures.* *Pesquisa Operacional,* v 26, p109-127. 2005
- DROGUETT, E.; GROEN, F.; MOSLEH, A.** *The Combined Use of Data and Experts Estimates in Population Variability Analysis.* *Reliability Engineering & System Safety,* v83, p311-321. 2003
- GREGORY, P.;** *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica™ Support.* Ed. Cambridge University Press. 2005
- HAWKSLEY, J.** *Some Social, Technical and Economic Aspects of the Risks of Large Plants,* CHEMRAWN III, 1984
- IEA - International Energy Agency.** *Key World Energy Statistics,* 2008. Disponível em <www.iea.org>
- JOHNSON, D.; WELKER, J.,** *Development of an Improved LNG Plant Failure Rate Data Base,* Applied Technology Corp., Norman, OK, 1981.
- KAPLAN, S.** *On a 'two-stage' Bayesian procedure for determining failure rates from experimental data.* *IEEE Trans Power Appar Syst,* p 195–199, 1983
- KELLY, D.; SMITH, C.** *Bayesian Inference for Probabilistic Risk Assessment : A Practitioner's Guidebook.* Ed Springer. London, 2011
- MME - Ministério de Minas e Energia.** *Balanco Energético Nacional,* 2011
- RAMOS, M.** *Desenvolvimento de Metodologia Bayesiana de Análise Quantitativa de Risco para Terminais Offshore de GNL: Aplicação para o Porto de Suape.* 121 folhas. Dissertação – Universidade Federal de Pernambuco. Recife, 2012
- SHAFAGHI, A.,** *Equipment Failure Rate Updating Bayesian Estimation,* the Mary Kay O'Connor Process Safety Center Symposium, p 260-268. 2006.
- SIU, N.; KELLY, D.** *Bayesian Parameter Estimation in Probabilistic Risk Assessment.* *Reliability Engineering & System Safety,* v 62, p 89-116, 1998
- The Netherlands Organization of Applied Scientific Research – TNO.** *CPR18E Guidelines for Quantitative Risk Assessment – Guideline for Quantitative Risk Assessment "Purple Book"* CPR18E. 3ed. The Hague, 2005.
- WIJNGAARDEN, W.; OOMEN, H.; HOORN, J.** *Offshore LNG Terminals: Sunk or Floated?* In: OFFSHORE TECHNOLOGY CONFERENCE. Houston, 2004
- WOODWARD, J.; PITBLADO, R.** *LNG Risk Based Safety – Modeling and Consequence Analysis.* 1 ed. Ed Wiley. New Jersey, 2010
- YUN, G.** *Bayesian-LOPA Methodology for Risk Assessment of an LNG Importation Terminal.* 288 folhas. Dissertação. Texas A&M University. 2007
- YUN, G.; ROGERS, W.; MANNAN, M.** *Risk Assessment of LNG Importation Terminals Using the Bayesian-LOPA Methodology.* *Journal of Loss Prevention in the Process Industries,* v 22, p91-96. 2009