

## 2 Referencial teórico

### 2.1. Energia eólica

O ar em movimento produz energia cinética. Esta energia cinética é chamada de energia eólica. A potência produzida pelo vento pode ser calculada através da derivada da fórmula da física da energia cinética ( $E_c = \frac{1}{2}mv^2$ ), que após alguma álgebra, resulta em:

$$P = \frac{1}{2}\rho Av^3 \quad (1)$$

Onde  $P$  é a potência disponível do vento (em watts),  $\rho$  é a massa específica do ar (ou densidade absoluta, em  $\text{kg/m}^3$ ),  $A$  é a área de secção transversal do cilindro que é ultrapassada pelo vento ( $\text{m}^2$ )<sup>1</sup> e  $v$  é a velocidade do vento em metros por segundo.

A potência é altamente sensível à velocidade do vento, um aumento de apenas 10% da velocidade do vento, gera 33% a mais de potência disponível (PINTO, 2012). Conclusão tão importante quanto essa, também em (PINTO, 2012), é a de que regimes com mesma velocidade média de vento não resultam em mesma quantidade de energia gerada. Variações na velocidade do vento ao longo de um período  $t$  de tempo tendem a ser positivas na geração de potência para uma mesma velocidade média.

---

<sup>1</sup> $(A = \frac{\pi D^2}{4})$ , com  $D$  sendo o diâmetro do rotor

O valor da densidade do ar,  $\rho$ , é de 1,225 kg/m<sup>3</sup> em condições climáticas consideradas normais, ou seja, ao nível do mar e temperatura ambiente de 15° Celsius. Para o caso brasileiro, normalmente se fazem necessárias correções a este valor (AMARANTE *et al.*, 2001). Estas correções podem ser feitas através de fórmulas (PINHO *et al.*, 2008) conhecendo-se a pressão atmosférica ( $P_y$ , em hPa) e a temperatura ( $T$ , em Celsius):

$$\rho = \frac{0,3484P_y}{273,185 + T} \quad (2)$$

Também se pode utilizar uma fórmula que necessita apenas da altitude ( $z$ , em metros), porém menos precisa:

$$\rho = 1,225 - (1,194 \times 10^{-4})z \quad (3)$$

A fórmula para calcular a potência apresentada inicialmente é apenas conceitual, em Pinho (2008) encontra-se referência à lei (ou limite) de Betz:

$$P = \frac{1}{2} C_p \rho A v^3 \quad (4)$$

Onde  $C_p$  representa o coeficiente de potência, parcela máxima do vento que pode ser efetivamente aproveitada pela turbina eólica e cujo máximo teórico é definido pelo limite de Betz ( $16/27 \approx 0,593$ ). Este valor é um máximo teórico para o aproveitamento de potência, porém os resultados práticos são ainda menores, chegando a 80% do limite nas melhores condições possíveis, ou aproximadamente 47% de eficiência na transformação da potência de vento em potência de um eixo de geração (PINTO, 2012). As fontes de perda mais comuns encontradas são as de perda aerodinâmica nas pás, perdas nos componentes da máquina e perdas relacionadas ao fato de o rotor eólico operar apenas em uma determinada faixa de velocidades de vento.

A potência extraível de um aerogerador varia com o cubo da velocidade do vento, tornando este parâmetro o mais significativo no momento de transformar a energia cinética em energia mecânica de rotação (PINHO *et al.*, 2008). Alguns parâmetros relacionados à velocidade de vento merecem ser mencionados:

- $V_p$ : Velocidade de partida (*cut-in*) – velocidade de vento mínima para que o rotor saia de seu estado de repouso inicial, produzindo torque mecânico e iniciando o movimento das pás e consequente geração de energia
- $V_g$ : Velocidade mínima de geração – valor mínimo de velocidade de vento na qual a geração de energia ainda é possível, determinada pelas perdas na transmissão. É normalmente um valor de velocidade levemente superior ao de partida
- $V_n$ : Velocidade nominal – velocidade de vento na qual a potência nominal (PN) do aerogerador é extraída;
- $V_c$ : Velocidade de corte (*cut-out*) – valor de velocidade de vento na qual a regulação atua para interromper o movimento do rotor eólico, protegendo-o contra cargas excessivamente altas e danos estruturais.

Pode haver, ainda, um quinto valor de velocidade de vento, pouco utilizado, denominado velocidade de controle, na qual o rotor sofre a ação de uma regulação (controle por estol) de modo a reduzir sua potência de saída.

O fator que se mostra mais relevante, como dito anteriormente, para o cálculo da potência gerada pelo vento é a velocidade, por estar ao cubo na equação ((1) apresentada. Um aumento ou redução na velocidade do vento será, portanto, multiplicado. Segundo (PINHO *et al.*, 2008), a velocidade do vento utilizada em projetos eólicos é resultante do cálculo de uma velocidade média em um determinado período de tempo, sendo definida como:

$$V_m = \frac{1}{N} \sum_{i=1}^N V_i \quad (5)$$

Ainda em (PINHO *et al.*, 2008), a maneira mais usual de se analisar graficamente dados estatísticos de vento é através dos dados médios da velocidade. Uma maneira menos usual, porém relevante, é a distribuição horária média da velocidade do vento, também conhecida como dia médio ou dia típico e que corresponde a uma velocidade diária média ao longo do dia. A aproximação por uma variável contínua é válida e extensamente apresentada na literatura, inclusive em (PINHO *et al.*, 2008), (PINTO, 2012) e (AMARANTE *et al.*, 2001).

A distribuição da velocidade do vento pode ser representada por uma função densidade de probabilidade  $f(v)$  como descrito na literatura:

$$V_m = \int_0^{\infty} v \times f(v) dv \quad (6)$$

Esta função  $f(v)$  é a função densidade de probabilidade da velocidade do vento. A distribuição utilizada pelos autores supracitados para a função  $f(v)$  é a distribuição de probabilidade de Weibull definida como:

$$f(v) = \frac{k}{C} \left(\frac{v}{C}\right)^{k-1} e^{-\left(\frac{v}{C}\right)^k} \quad (7)$$

onde  $v$  é a velocidade do vento (m/s),  $k$  é um fator de forma (m/s) e  $C$  é o fator de escala (adimensional). Os valores de  $k$  costumam variar entre 2 e 3. Em determinadas épocas do ano, no Nordeste brasileiro, o fator pode atingir valores superiores a 6. Estes porém são casos excepcionais.

Os fatores  $k$  e  $C$  são parâmetros e podem ser calculados aproximadamente por (PINHO *et al.*, 2008):

$$k = \left(\frac{\sigma}{V_m}\right)^{1,086}, \quad C = \frac{V_m}{\Gamma\left(1 + 1/k\right)} \quad (8)$$

Os parâmetros médios de  $k$  da distribuição de Weibull encontrados no Ceará foram de 3,44 no interior e 4,67 para o litoral (PINTO, 2012).

A direção do vento é outro fator que não pode ser negligenciado ao se analisar a energia eólica. Ela indica de onde o vento é proveniente. Utiliza-se a rosa dos ventos para representar dados anuais, sazonais ou mensais, sendo útil para determinar a posição das turbinas eólicas. Deseja-se o menor número de obstáculos possíveis na direção de onde provém o vento.

## 2.2. Tipos de modelos de previsão

Tendo o conhecimento de como se comporta a variável aleatória da velocidade do vento, o passo seguinte é conseguir um modelo de previsão que consiga captar ao máximo as variações nesta variável e que seja, por conseguinte, o mais preciso possível. Há diversos horizontes de previsão. Para Jung (JUNG; BROADWATER, 2014), as previsões podem ser divididas em quatro horizontes

diferentes de tempo, são eles: o muito curto prazo (poucos segundos à frente até trinta minutos), o curto prazo (entre 30 minutos à frente e seis horas), o médio prazo (entre seis e vinte e quatro horas) e o longo prazo (entre um e sete dias). Previsões de muito curto prazo são usadas para o controle da turbina e rastreamento de carga. Previsões de curto prazo são utilizadas para o compartilhamento de pré-carga. As de médio prazo para gestão do sistema de energia e comercialização de energia e as de longo prazo para a programação e manutenção das turbinas eólicas e otimização dos custos de operação.

### **2.2.1. Modelos de previsão físicos (determinísticos)**

Uma das técnicas de previsão da velocidade do vento é aquela baseada em dados físicos descrita por Kariniotakis (KARINIOTAKIS *et al.*, 2004) e Focken (FOCKEN; LANGE, 2006). Este tipo de previsão necessita de um grande conhecimento das características físicas do parque eólico e de seus arredores, indo desde a curvatura do hélice aos possíveis obstáculos do terreno e às condições atmosféricas da região. Apesar de dependerem de um grande número de dados diferentes, os métodos físicos possuem a vantagem de não precisarem de séries históricas de velocidade do vento. Por outro lado, este tipo de previsão está sujeito a uma variedade de tipos de erros de medida pois lida com diferentes fontes de dados. Os métodos físicos são, em geral, superiores aos métodos estatísticos no curtíssimo prazo, pois os últimos falham em capturar a dinâmica da atmosfera e outras condições que podem levar a pequenas variações na velocidade do vento no curto prazo.

Kariniotakis (KARINIOTAKIS *et al.*, 2004), Costa (COSTA *et al.*, 2008) e Jung (JUNG; BROADWATER, 2014) oferecem uma descrição dos vários métodos de previsão baseada em dados físicos. Dentre os modelos mais conhecidos estão o *Prediktor*, desenvolvido na Dinamarca, o *Previento*, desenvolvido na Alemanha, o *Local Pred*, desenvolvido na Espanha e o *eWind* desenvolvido nos Estados Unidos.

### 2.2.2. Modelos de previsão estatísticos (estocásticos)

Os modelos estatísticos mais simples, por outro lado, se baseiam na metodologia de Box e Jenkins (BOX; JENKINS, 1976). Estes modelos precisam de longas séries históricas de vento para serem ajustados, entretanto são menos custosos do que os modelos físicos que necessitam de uma coleção de informações de medição complexa. De forma geral, esta metodologia pode ser descrita da seguinte forma:

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + c + \epsilon_t \quad (9)$$

onde  $y_t$  representa a velocidade do vento no tempo  $t$ ,  $\phi_i$  é o coeficiente auto regressivo e  $\theta_j$  é o coeficiente de média móvel.  $\epsilon_t$  é um termo de erro aleatório e  $c$  uma constante. O modelo descrito pela equação ((9) é um  $ARMA(p, q)$ <sup>2</sup> e deve ser aplicado apenas para séries estacionárias. Caso as séries não sejam estacionárias, elas devem ser diferenciadas  $d$  vezes até que a estacionariedade seja alcançada, o modelo será então um  $ARIMA(p, d, q)$ <sup>3</sup>. Outra forma de tratar a não estacionariedade da série de vento consiste em transformar a série de vento em uma série gaussiana e depois padronizá-la para que tenha média igual a zero e variância igual a um (ZHANG *et al.*, 2011). Para determinar os valores de  $p$  e  $q$  deve-se observar a função de autocorrelação e a função de autocorrelação parcial das séries de vento para verificar em quantos períodos existe uma correlação temporal dos ventos.

Os modelos da família ARIMA são populares pela sua fácil aplicação e bons resultados de previsão, além de permitirem inferência e testes de hipótese, ao contrário dos modelos físicos. Este é o modelo estatístico mais simples conforme descrito por Shiyan (LEI *et al.*, 2009), Foley (FOLEY *et al.*, 2012), Dongmei (DONGMEI *et al.*, 2011) e Jung (JUNG; BROADWATER, 2014). A interpretação da equação ((9) também é bem simples, o vento no período  $t$  depende apenas do vento nos períodos passados e dos termos de erro passados, entretanto existem algumas condições nos valores dos parâmetros para que o

<sup>2</sup> Autoregressive Moving Average

<sup>3</sup> Autoregressive integrated moving average

modelo seja estável. É possível incluir outros regressores ao modelo, contanto que estes sejam estacionários e atendam certas condições, caso contrário o problema se torna muito mais complexo e requer outras metodologias como o vetor autorregressivo (VAR). Os modelos estatísticos convencionais AR, MA, ARMA, abordagem Box-Jenkins e SARIMA visam prever o muito curto prazo e o curto prazo, na maioria dos casos, segundo Jung (JUNG; BROADWATER, 2014). Quanto maior o horizonte de previsão, maior o erro de previsão e o intervalo possível para os valores previstos.

Os modelos ARIMA são muito simples e em alguns casos podem não reter informação suficiente para uma boa previsão. Por este motivo outros modelos são acoplados à forma tradicional do ARMA da equação ((9) em determinadas ocasiões. Os modelos mais conhecidos entre estes são aqueles da família ARCH<sup>4</sup>. Estes modelos são aplicados ao termo de erro da equação ((9) e visam modelar a heterocedasticidade<sup>5</sup> condicional do erro. Em outras palavras, sabe-se que a variância do erro (parte estocástica do modelo) não é constante no tempo e modelos ARCH, GARCH<sup>6</sup> e EGARCH<sup>7</sup> visam modelar este fenômeno.

Em geral, os modelos da família ARCH são utilizados em conjunto com modelo ARMA. A intuição por trás do modelo é criar uma estrutura para a variância do erro da equação do ARMA definido como  $\epsilon_t$ :

$$\epsilon_t = \sigma_t z_t \quad (10)$$

onde  $z_t$  representa um ruído branco e a variância  $\sigma_t^2$  é modelada da seguinte forma:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 \quad (11)$$

onde  $\alpha_0$  é uma constante e  $\alpha_i$  é o parâmetro de cada defasagem do termo de erro elevado ao quadrado.

---

<sup>4</sup>Autoregressive conditional heteroskedasticity

<sup>5</sup> De forma bem simples, o termo heterocedasticidade significa que a variância condicional ao tempo do termo de erro não é constante, ou seja, ela varia com o tempo.

<sup>6</sup>generalized autoregressive conditional heteroskedasticity

<sup>7</sup> exponential general autoregressive conditional heteroskedastic

Um dos problemas do modelo ARCH na forma apresentada acima é a necessidade de um grande número de parâmetros para eliminar a autocorrelação dos erros ao quadrado. Para tornar o modelo mais parcimonioso existe uma forma generalizada do ARCH conhecida como GARCH, inicialmente proposto por (BOLLERSLEV, 1986). A principal diferença entre os dois modelos está na forma de modelar o termo  $\sigma_t^2$ , tendo o GARCH a seguinte expressão:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 \quad (12)$$

onde  $\alpha_0$  é uma constante,  $\alpha_i$  é o parâmetro de cada defasagem do termo de erro elevado ao quadrado e  $\beta_i$  é o parâmetro de cada defasagem da variância do termo de erro. O modelo acima definido é um *GARCH(p, q)*.

Trabalhos com aplicação destes modelos em séries de vento podem ser vistos em (LV; YUE, 2011) e (WANG *et al.*, 2012). Os dois trabalhos mostraram que as previsões de curto prazo obtiveram menores erros quando os modelos são controlados para heterocedasticidade condicional via modelos ARCH.

### 2.2.3.

#### Modelos de redes neurais e modelos híbridos

Um outro método estatístico é aquele conhecido como *Artificial Neural Network Approach* (ANN). O ANN consiste em utilizar previsões de outros modelos e variáveis meteorológicas que são transformadas em previsões de potência ou velocidade do vento. Uma característica dos modelos de redes neurais é o auto aprendizado de acordo com novas informações, entretanto o ANN requer grande quantidade de dados para ser bem calibrado e obter bons resultados. Kani e Riahy (2008) utilizaram o ANN com cadeia de Markow para previsões de curtíssimo prazo, sendo as cadeias de Markow utilizadas para modificar as primeiras previsões do ANN de acordo com padrões de longo prazo nas variáveis. Os resultados mostraram que o método é bastante eficiente. Jursa (2007) também obtém bons resultados com a utilização do ANN para reduzir erros de previsão.

Amjady (AMJADY *et al.*, 2011) utiliza um método baseado em redes neurais de Ridgelet (RNN). Para definir as informações de entrada no modelo é utilizado o critério de informação mútua no conjunto de possíveis variáveis relevantes para o modelo. Os resultados também foram bons.

Uma forma interessante de se prever o comportamento dos ventos é utilizar tanto sua velocidade quanto sua direção como dados. Esta é a proposta das redes neurais conhecidas como *Complex-Valued Neural Network* (CVNN) apresentadas por Kitajima e Yasuno (KITAJIMA; YASUNO, 2010).

Os métodos de redes neurais apresentados acima são, em geral, modelos híbridos, ou seja, utilizam modelos estatísticos e físicos ao mesmo tempo. O objetivo é utilizar o máximo da informação disponível e buscar explorar as vantagens das previsões estatísticas e dos dados físicos. As vantagens de se combinar modelos para diminuir o risco de previsões ruins são discutidas por Hibon (HIBON; EVGENIOU, 2005).

Além dos modelos apresentados acima, existem também aqueles que utilizam lógica *fuzzy*. Eles consistem em mapear os dados de entrada usando valores numéricos e linguísticos como variáveis categóricas (rápido-lento, quente-frio, alto-médio-baixo). Métodos *fuzzy* são utilizados diante da dificuldade de se medir um conjunto de variáveis relevantes com precisão. Aplicações com resultados relevantes dessa metodologia podem ser vistos em Sideratos (SIDERATOS; HATZIARGYRIOU, 2007) e Hong (HONG *et al.*, 2010).

Os modelos de Redes Neurais são muito flexíveis e a decisão de qual modelo utilizar pode ser um problema. Muitas vezes estas técnicas são muito específicas e requerem grandes quantidades de dados variados para a aprendizagem do modelo. Além disso, não existe qualquer tipo de teoria por trás que garanta o funcionamento dos modelos, ou seja, são modelos puramente computacionais.

### **2.3. Otimização**

#### **2.3.1. Problemas de otimização**

*Otimização* é a “ciência do melhor” no sentido em que esta ajuda não somente a efetuar uma decisão razoável, mas a melhor decisão possível dado um conjunto de restrições que insere o problema dentro de um universo de possibilidades limitadas. Modelos matemáticos de programação proveem o *framework* apropriado para se abordar estas decisões de otimização de forma precisa e formal (CONEJO *et al.*, 2006).

O ‘alvo’ ou objetivo a ser maximizado (ou minimizado) é expresso em forma de função matemática denominada de “função objetivo” por ser esta a função que se deseja encontrar o valor máximo (ou mínimo).

As restrições devem ser satisfeitas e definem o que se chama de “região de viabilidade” do problema. Importante notar que esta região deve contar mais de uma solução possível, caso contrário o problema perde interesse prático. As restrições são expressas por funções de valores reais das variáveis de decisão, sempre através de equações ou inequações (CONEJO *et al.*, 2006)..

Um problema de otimização matemática adota a seguinte forma:

$$\begin{aligned} & \text{Minimize } f_0(x) \\ & \text{sujeito a: } f_i(x) \leq b_i, \quad i = 1, \dots, m. \end{aligned} \quad (13)$$

Neste caso, o vetor  $x = (x_1, \dots, x_n)$  é a variável de otimização do problema, a função  $f_0: \mathbb{R}^n \rightarrow \mathbb{R}$  é a função objetivo, as funções  $f_i: \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \dots, m$  são as funções de restrição e as constantes  $b_1, \dots, b_m$  são os limites ou fronteiras para as restrições. Um vetor  $x$  é chamado ótimo ou uma solução do problema acima se tiver o menor valor para a função objetivo dentre todos os vetores que satisfaçam as restrições: para cada  $z$  com  $f_1(z) \leq b_1, \dots, f_m(z) \leq b_m$  temos  $f_0(z) \geq f_0(x^*)$  (BOYD; VANDENBERGHE, 2004).

Este problema da equação ((13) se caracteriza como sendo um problema de programação linear (BOYD; VANDENBERGHE, 2004) se tanto a função objetivo quanto as restrições são lineares, id est, satisfazem a seguinte propriedade:

$$f_i(\alpha x + \beta y) = \alpha f_i(x) + \beta f_i(y) \quad (14)$$

O problema de otimização apresentado é uma abstração do problema de escolher o melhor vetor possível no  $\mathbb{R}^n$  de um conjunto de vetores possíveis. A variável  $x$  representa a escolha feita; a restrição  $f_i(x) \leq b_i$  representa requerimentos da empresa ou limitações que alteram as escolhas possíveis, e a função objetivo  $f_0(x)$  representa o custo de se escolher  $x$  (também pode ser pensada como a utilidade em se escolher  $x$ ). A solução da otimização corresponde à escolha que possui o menor custo (ou maior utilidade), e que satisfaça os requerimentos da firma.

Em problemas de otimização de portfólio, o que se busca é a melhor forma de se investir capital em um conjunto de  $n$  ativos. A variável  $x_i$  representa o investimento no  $i$ -ésimo ativo, então o vetor  $x \in \mathbb{R}^n$  descreve a alocação global do portfólio através do conjunto dos ativos. As restrições podem representar limite de orçamento, a necessidade de que os investimentos sejam não-negativos (sem vendas a descoberto ou alavancagens) e um valor médio mínimo aceitável de rendimento para o portfólio como um todo. A função objetivo pode representar uma medida de risco global ou a variância do retorno do portfólio. No caso apresentado acima, o problema de otimização corresponde a escolher a alocação de portfólio que minimiza o risco da carteira, entre todas as possíveis alocações que atendem os requisitos da empresa (BOYD; VANDENBERGHE, 2004).

### 2.3.2. Programação Linear (LP)

Um problema genérico encontrado em programação linear pode ser formalizado matematicamente da seguinte forma:

$$\begin{aligned} & \text{Minimize } c^T x + d \\ & \text{sujeito a:} \\ & Gx \leq h \\ & Ax = b \end{aligned} \tag{15}$$

Onde  $G \in \mathbb{R}^{m \times n}$  e  $A \in \mathbb{R}^{p \times n}$ . Funções LP são problemas convexos também. Na equação da função objetivo (15) é comum que se omita a constante  $d$  pois esta não afeta o conjunto ótimo ou possível de soluções. Como é possível maximizar uma função objetivo afim  $c^T x + d$  minimizando  $-c^T x - d$  (que também é convexa), também se pode referir a um problema de maximização com objetivo afim e restrições como uma LP (BOYD; VANDENBERGHE, 2004).

Na Figura 3 pode-se observar uma representação geométrica de um problema de otimização linear. O conjunto de soluções factíveis é um poliedro  $P$ . A função objetivo  $c^T x$  é linear, portanto suas curvas de nível são hiperplanos ortogonais a  $c$  (as linhas pontilhadas). O ponto  $x^*$  é o ponto ótimo; este se encontra dentro do poliedro e o mais distante na direção de  $-c$ .

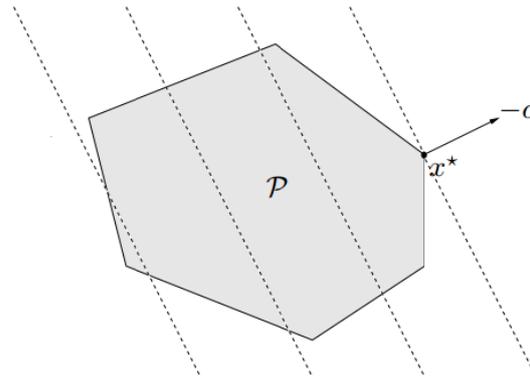


Figura 3 – Solução de problema de otimização  
 Fonte: Boyd e Vandenberghe, 2004

Um algoritmo é chamado de **método de solução** para uma classe de problemas de otimização quando ele é capaz de computar uma solução, dado uma classe particular de problema, *i.e.*, uma instância do problema. (BOYD; VANDENBERGHE, 2004). Mesmo quando a função objetivo e as restrições são regulares o problema de otimização pode ser surpreendentemente difícil de se resolver. Abordagens para o problema envolvem, normalmente, um grande comprometimento de tempo computacional e podem terminar sem encontrar solução.

Algumas exceções importantes são problemas de mínimos-quadrado e de programação linear. Há algoritmos efetivos e confiáveis disponíveis para a solução destes tipos de problema. Não há soluções com fórmulas analíticas para a programação linear como ocorre nos problemas de mínimos-quadrados, porém há vários métodos conhecidos e eficientes para solucioná-los. É o caso do algoritmo Simplex e alguns algoritmos mais recentes de ponto de interior. Não é possível afirmar quantas operações serão necessárias para a solução de um problema LP, porém é possível determinar um limite. A complexidade, na prática, é da ordem de  $O(n^2m)$ <sup>8</sup> (assumindo  $m \geq n$ ). A solução para problemas com centenas de variáveis e milhares de restrições pode ser conseguida em alguns segundos em computadores pessoais (BOYD; VANDENBERGHE, 2004).

<sup>8</sup> Sendo  $f(n)$  e  $g(n)$  duas funções positivas.  $f = O(g)$  (ou seja,  $f$  não cresce mais rápido que  $g$ ) se existir uma constante  $c > 0$  tal que  $f(n) \leq c \times g(n)$ .

### 2.3.3. Value at Risk –VaR

O *Value at Risk* (VaR) é definido como sendo uma medida estatística que permite àquele que o utiliza medir o risco inerente de uma determinada carteira (HORTA, 2000). O detentor de uma carteira de ativos ou títulos, independentemente de ser um pequeno investidor ou um investidor institucional, sempre estará interessado no nível de risco ao qual sua carteira está exposta. Quatro questões são imprescindíveis de serem respondidas para qualquer gestor e o cálculo do VaR de uma carteira é um instrumento capaz de apresentar uma estimativa, com um determinado nível de confiança, que responde a estes questionamentos. São eles: Quanto dinheiro poderei perder com a minha carteira atual caso a mantenha por mais um dia? Se eu comprar ações de uma empresa aumentarei o risco global da minha carteira? Se sim, qual é o acréscimo desse risco em termos de montante? E como deverei recompor a minha carteira se só estiver, por exemplo, disposto a perder no máximo 5% do seu valor? (HORTA, 2000)

Mas o que é exatamente o VaR? De forma geral, Damodaran (DAMODARAN, 2007) determina o VaR como sendo a medida do potencial de perda de valor de um ativo de risco ou portfólio de risco durante um período definido de tempo para um determinado intervalo de confiança. E exemplifica: se o *Value at Risk* em um ativo é de 100 milhões, em uma semana, com 95% de confiança, existe apenas 5% de chance de que o valor deste ativo caia mais de 100 milhões durante o período de uma semana.

Damodaran identifica três formas de se calcular o VaR (DAMODARAN, 2007), o método da Variância-Covariância que se baseia na derivação da densidade de probabilidade dos valores potenciais que podem ser atingidos e a utiliza para calcular o VaR; o método das simulações históricas, que consiste em estimar um portfólio criando uma série temporal hipotética dos retornos deste portfólio ao longo do tempo, obtidos executando-o através dos dados históricos reais e computando o retorno que teria ocorrido em cada período. Por último, o método de simulação de Monte Carlo. Neste método são especificadas densidades de probabilidade para cada fator de risco do mercado e como estes fatores se movem conjuntamente. Este último método é o que apresenta maior flexibilidade para se escolher a distribuição dos retornos e para que seja possível efetuar

julgamentos subjetivos e dados externos ao problema que se deseja resolver, porém é custoso computacionalmente.

#### **2.3.4. Conditional Value at Risk –CVaR**

O *Conditional Value at Risk* (CVaR) é a perda esperada condicional, sob a condição que se exceda o VaR. Em outras palavras, o CVaR responde ao seguinte questionamento: “Como espera-se incorrer em perdas quando a situação for indesejada?” (SALAHİ *et al.*, 2013).

O VaR é uma medida de risco muito criticada na literatura por não ter algumas propriedades consideradas essenciais, como por exemplo, a da subatividade, *id est*, que o risco de um portfólio pode ser maior do que a soma dos riscos de cada um dos ativos que compõem o portfólio (ACERBI; TASCHE, 2002). Outro problema apontado é que o VaR não leva em consideração a severidade de um evento ruim.

A construção do CVaR segue, portanto, uma linha que ‘corrija’ os defeitos do VaR sem perder suas muitas qualidades. O CVaR é calculado pegando-se a média ponderada das perdas possíveis condicionadas a que estas sejam iguais ou excedam o VaR especificado (XIONG; IDZOREK, 2011). Outros termos comuns para CVaR encontrados na literatura são *Expected Shortfall* (ES), *Mean Shortfall tail VaR* e *Expected tail loss*. Intuitivamente, o CVaR é uma medida mais completa de mensuração de risco do que o VaR e diversos estudos mostraram que o primeiro possui propriedades mais atraentes que o último.

Ao compararmos otimizações de média com restrição de CVaR (M-CVaR) a otimizações de média com restrição de variância (MVO, uma das mais simples e, por conseguinte, mais comumente utilizada) encontramos que se o retorno das classes dos ativos é normalmente distribuído ou se o método utilizado para estimar o CVaR considera somente os dois primeiros momentos estatísticos, tanto a otimização MVO quanto a M-CVaR levam à mesma alocação de ativos. Porém, quando há valores divergentes de assimetria e curtose entre os ativos, os resultados se mostram bastante diferentes. Mais especificamente, as otimizações de M-CVaR apresentam preferência por ativos com maior assimetria, menor curtose e menor variância (XIONG; IDZOREK, 2011).

Uma formalização possível e que exemplifica uma otimização M-CVaR é apresentada a seguir (SALAH *et al.*, 2013). Considere os ativos  $S_1, \dots, S_n$ ,  $n \geq 2$ , com retornos aleatórios. Suponha que  $\mu_i$  denote o retorno esperado do ativo  $S_i$ , e considere também  $x_i$  como a proporção detida do  $i$ -ésimo ativo. Pode-se representar o retorno esperado do portfólio  $x$  como a seguir:

$$E[x] = \mu_1 x_1 + \dots + \mu_n x_n = \mu^T x. \quad (16)$$

Além disso, assume-se que o conjunto de portfólios possíveis é um poliedro não vazio e pode ser representado como  $\Omega = \{x | Ax = b, Cx \geq d\}$  onde  $A$  é uma matriz  $m \times n$ ,  $b$  é um vetor  $m$ -dimensional,  $C$  é uma matriz  $p \times n$ , e  $d$  é um vetor  $p$ -dimensional. Em particular, uma das restrições no conjunto  $\Omega$  é  $\sum_{i=1}^n x_i = 1$ .

Deixe que  $f(x, y)$  denote a função de perda quando se escolhe o portfólio  $x$  de um conjunto de portfólios possíveis, e  $y$  a realização dos eventos aleatórios (o vetor dos retornos dos ativos de  $n$  ativos). Considere que a perda nos retornos do portfólio,  $f(x, y)$ , seja uma função convexa (linear) igual a:

$$f(x, y) = -y^T x = -[y_1 x_1 + \dots + y_n x_n] \quad (17)$$

Assume-se que o vetor aleatório  $y$  tenha uma função densidade de probabilidade denotada por  $p(y)$ . Para um vetor de decisão  $x$ , a função de distribuição acumulada da perda associada a esse vetor é computada da seguinte forma:

$$\psi(x, y) = \int_{f(x, y) \leq \gamma} p(y) dy. \quad (18)$$

Em seguida, para um dado nível de confiança  $\beta^9$ , o  $\beta$ -VaR associado com o portfólio  $x$  é representado como:

$$VaR_\beta(x) = \min\{\gamma \in \mathbb{R} | \psi(x, \gamma) \geq \beta\}. \quad (19)$$

---

<sup>9</sup> $\beta$  é o nível de confiança, em alguns momentos o parâmetro  $\alpha$  (nível de significância) irá aparecer no lugar de  $\beta$ , sendo  $\alpha = 1 - \beta$ .

De forma similar, define-se o  $\beta$ -CVaR associado ao portfólio  $x$  como:

$$CVaR_{\beta}(x) = \frac{1}{(1 - \beta)} \int_{f(x,y) \geq VaR_{\beta}(x)} f(x,y)p(y) dy. \quad (20)$$

Em sua forma discretizada e simplificada, proposta pela primeira vez por Rockafellar (ROCKAFELLAR; URYASEV, 2000) e que se tornou muito popular:

$$CVaR_{\beta}(x) = z + \frac{1}{q(1 - \beta)} \sum_{k=1}^q u_k$$

*sujeito a:* (21)

$$u_k \geq 0 \text{ e } ex^T y_k + \beta + u_k \geq 0 \text{ para } k = 1, \dots, r.$$

Onde  $q$  representa o número de cenários,  $\beta$  o nível de confiança,  $z$  é o VaR e  $u_k$  representa uma variável auxiliar para computar apenas os cenários ‘ruins’.