

2 Pagamentos por Capacidade

2.1 Introdução

O custo do fornecimento da energia elétrica pode ser dividido em custos da demanda, ou da capacidade de geração e custos da energia, relacionados direta e respectivamente com os custos fixos e os custos variáveis.

Os pagamentos pela capacidade de geração encontram sua origem na teoria da fixação de preços da carga de ponta, introduzida por [Boiteaux, 1960; Steiner, 1957], entre outros. Segundo esta teoria (*peak-load pricing*), supondo-se um período de ponta e outro fora da ponta, com a mesma capacidade disponível, à energia deve ser dado preço diferenciado em cada período, segundo os custos marginais de operação. Deve ser criado também um pagamento pela capacidade, com o fim de recuperar os custos fixos, ou custos da capacidade instalada, a ser cobrado apenas dos consumidores com demanda máxima coincidente com a ponta de demanda do sistema.

Este enfoque tradicional criou entretanto uma situação injusta, pois os consumidores fora de ponta, como também necessitam de capacidade disponível, utilizavam a capacidade 'adquirida' pelos consumidores de ponta, numa espécie de subsídio cruzado. Estes consumidores chegaram a se acharem no direito de reclamar a 'propriedade' da capacidade de geração instalada, uma vez que apenas eles pagavam pelo seu custo fixo, passando a exigir dos consumidores fora de ponta um pagamento pela capacidade que agora lhes pertencia. Diversas soluções foram propostas, para transferir parte dos custos de capacidade para os consumidores fora de ponta, por exemplo, por [Weintraub, 1970].

Uma justificativa técnica da necessidade dos consumidores fora da ponta também assumirem custos de capacidade segue. O consumo de energia cresce ao longo do tempo, devido ao aumento do número de consumidores e de equipamentos elétricos em uso, bem como o aumento da intensidade de seu uso. Usinas adicionais são portanto necessárias para atender, não só aos aumentos da carga de ponta mas também à expansão da carga de base. Desta forma, os custos de capacidade adicionais, envolvidos na instalação de novas usinas de base, precisam ser suportados por todos os consumidores, independentemente do período de uso, uma vez que estas usinas estão continuamente em uso, em ambos os períodos.

O enfoque tradicional só é realmente válido em casos de sistemas com geração homogênea, ou seja, onde aparece uma única tecnologia de geração. Com tecnologia heterogênea, fator importante no alcance da eficiência nos preços da energia elétrica, os custos marginais fora de ponta também devem incluir uma parte de custos de capacidade. Esta heterogeneidade gera um outro problema, que também tem a ver com capacidade de geração, descrito em seguida.

Em sistemas de energia elétrica com este perfil e com grande predominância de tecnologias térmicas, as usinas de mais alto custo de produção, com certeza serão menos despachadas do que usinas mais 'baratas', criando-se uma alta volatilidade na sua remuneração³. Elas podem ficar até anos sem gerar, caso ocorra um período longo de grande precipitação pluviométrica (baixo custo marginal de curto prazo). Usinas sob tal condição, conhecidas como usinas de ponta, poderiam ser alvo de desativação por seus proprietários, em função do prejuízo causado pela manutenção dos seus custos fixos, sem faturamento para cobri-los. Isto viria a diminuir a disponibilidade de geração e, em consequência, a confiabilidade do sistema, a longo prazo. Em sistemas onde predominam as plantas hidrelétricas, como é o caso brasileiro, de térmica, mas com menor intensidade e frequência.

Conclui-se assim que a solução deste problemas e os pagamentos pela capacidade estão intimamente relacionados, através da cobertura dos custos fixos. Segundo [Oren, 2000b], algumas justificativas para se implantarem os pagamentos por capacidade seriam:

- definição de preços para a carga de ponta;
- definição de preços e remuneração para a confiabilidade;
- recuperação de custos.

Caso se decida criar algum tipo de remuneração específica para compensar a capacidade ociosa das usinas de ponta, deve haver a preocupação de que ela seja cobrada dos consumidores corretos, e paga aos agentes corretos.

³ Em sistemas onde predominam as plantas hidrelétricas, como é o caso brasileiro, as usinas de ponta geralmente não são usinas hidrelétricas, pois o custo marginal de operação do sistema é mais influenciado pelo valor atual do custo futuro do uso da água do que pelo custo de operação das usinas térmicas.

2.2 Pagamentos pela Capacidade e Confiabilidade a Longo Prazo

Em ambientes de mercado observa-se uma certa negligência em relação à política de confiabilidade de longo prazo e investimento em capacidade, o que veio a produzir mecanismos desastrosos de preços em alguns sistemas. De um modo geral, os mercados não conseguem, por si sós, definir tais políticas, exigindo-se assim uma interferência regulatória para a definição de uma combinação de energia em tempo real, reservas operativas e capacidade instalada, respaldadas por uma política de preços, também com algum caráter regulatório. Sem isso, os mercados atuais de energia, por seus problemas naturais, tais como a elasticidade de demanda muito baixa, tendem a sub-investir em geração. Se não existir uma política de confiabilidade que garanta preços suficientemente altos, pelo menos durante um certo número de horas por ano, os geradores não terão seus custos fixos cobertos, e assim não investirão em novas plantas.

Os picos de preço da energia, caso ocorram, e os preços da capacidade instalada, caso sejam praticados, proporcionam sinais econômicos que induzem o investimento de longo prazo em geração, por proporcionarem a cobertura dos custos fixos. Como consequência ocorre um incremento da confiabilidade, o que faz com que os picos de preço da energia e os preços da capacidade instalada devam ser os principais alvos da regulação [Stoft, 2002].

Assim, a confiabilidade a longo prazo de um sistema elétrico, pode ser considerada como dependente da recuperação dos custos fixos (custos de capacidade). Com isto, podem ser caracterizadas diversas formas para se incrementar esta confiabilidade.

Nos sistemas onde exista uma estrutura de mercado competitivo, no qual os geradores ofertam apenas energia, com base nos seus custos marginais, e o preço pago a todos os geradores despachados é o preço de liquidação do mercado, a recuperação dos custos fixos viria através da diferença entre este preço de liquidação e os custos marginais de cada gerador, isto é, o seu lucro de curto prazo. A chamada 'função de lucro', relaciona o lucro esperado na venda de energia com a capacidade instalada do sistema, de forma que, quando esta se encontra baixa, com maior probabilidade de cortes de carga, os preços de energia são mais altos e o lucro maior. No caso inverso, as perdas de carga seriam menos freqüentes, e os preços raramente altos, o que levaria a um menor lucro [Stoft, 2002]. Estas duas situações se alternam, até que seja atingido um nível de equilíbrio para a capacidade instalada, tal que o lucro da venda da energia poderia cobrir

os custos fixos de uma nova unidade de ponta, que é de aproximadamente US\$6.00/MWh, em média [Stoft, 2002].

Este é o caso, por exemplo, do Nordpool, Austrália, Nova Zelândia e Califórnia. A necessidade de se incorporar nova geração ao sistema é definida pelo operador do sistema ou por um órgão regulador. Além disto, se existir aquisição em separado de serviços ancilares pelo operador, os geradores podem auferir rendimentos extras vendendo estes serviços, seja por meio de mercados de curto prazo ou contratos de longo prazo.

Uma outra maneira de se melhorar a confiabilidade de longo prazo, ou 'adequação da geração' em um sistema elétrico, é a especificação de requisitos de reservas de planejamento (capacidade) pelo operador do sistema ou por um órgão regulador. Nesse caso, as empresas distribuidoras são obrigadas a manter uma certa capacidade de geração, que inclui um percentual de reserva, acima da sua demanda de ponta, possível de ser disponibilizada dentro de um certo período de tempo. Esta reserva pode advir de capacitação própria ou ser adquirida de empresas geradoras, o que pode trazer o estabelecimento de mercados de capacidade específicos, formais ou informais, que dariam aos geradores a chance de ter receitas adicionais, relativas à sua capacidade de geração não utilizada. Neste caso, a função de lucro é muito diferente, pois quando a capacidade instalada é menor que o requisito definido, paga-se por ela um determinado valor a todos os geradores. Quando a capacidade instalada encontra-se acima do requisito, este pagamento não existe. Este procedimento é utilizado principalmente nos Estados Unidos, nos sistemas Pennsylvania/New Jersey/Maryland Interconnection (PJM), New York e New England.

Quando se adota a política de definição de requisitos de capacidade, a soma das duas funções de lucro, da energia e da capacidade, gera uma função de lucro total da qual se extrai um valor final de equilíbrio para a capacidade instalada. Este valor nem sempre representa o valor ótimo, o qual deve ser tal que o lucro de uma unidade de ponta seja superior a US\$ 6.00/MWh aproximadamente [Stoft, 2002]. Para se alcançar o valor ótimo, várias políticas podem ser adotadas:

- Definição de um limite de preço para a energia, quando há corte de carga.

Na ocorrência de corte de carga, o preço da energia é feito igual ao custo imposto pela interrupção involuntária do fornecimento, o VLL (*value of lost load*), conhecido no Brasil como custo unitário da interrupção, C_{int} . Este procedimento é conhecido como *VLL pricing*, e levaria a um nível ótimo de investimento em geração e a um nível ótimo de confiabilidade. Em geral, esta política é feita para garantir por volta de 3 horas por ano com a energia a preço igual ao VLL. O valor ótimo da duração do corte de carga pode ser determinado como a razão entre o custo fixo da unidade de ponta e o VLL;

- Definição de um limite de preço para a energia, quando os requisitos de reserva operativa não são alcançados.

Segundo esta política, denominada *operating-reserve pricing* o valor ajustado para o preço da energia seria bem mais baixo, sendo esta providência tomada quando o nível de reservas operativas estivesse baixo. Apesar destas reservas serem necessárias para a solução de problemas de confiabilidade de prazo muito curto, a definição de seus preços tem influência também na confiabilidade de longo prazo, uma vez que também incentivam nova geração. As suas principais vantagens seriam:

- diminuição da volatilidade dos pagamentos às unidades despachadas apenas nos períodos de ponta de demanda;
 - diminuição do risco dos investidores;
 - diminuição do poder de mercado.
- modificação dos requisitos de potência operativa;
 - penalização, no caso de haver deficiência de capacidade.

Ainda em ambientes competitivos, com preços baseados em custos marginais, no que diz respeito às unidades de base que possuem custos fixos mais altos, a cobertura de tais custos, considerando-se valores médios, seria feita a longo prazo. Isto se dá desde que não haja excesso de capacidade instalada e a competição seja perfeita, ou seja, não hajam barreiras para a entrada ou saída de participantes. Não existe portanto uma garantia da cobertura dos custos fixos, e se isto acontecer durante um período longo, pode vir a inviabilizar a operação da usina. Esta situação pode ser vista também em

ambientes não competitivos, caso os preços sejam também baseados em custos marginais.

A adoção de pagamentos específicos pela capacidade é uma forma de se compensar as unidades geradoras que não tenham seus custos fixos completamente cobertos, além de ser capaz de aumentar a confiabilidade a longo prazo, pela possibilidade de incentivo direto à implantação de geração nova. Dependendo da forma que forem definidos, eles podem também auxiliar na solução da questão das unidades de ponta não despachadas, passíveis até de desativação. Diversos caminhos podem ser usados para a implementação de tais pagamentos, como os abordados na Seção 2.3.

2.3 Formas de Remuneração da Capacidade

2.3.1 Tarifas de Capacidade

São pagamentos a serem feitos aos geradores, calculados geralmente com base na confiabilidade do sistema, conhecidos pelo lado da carga como 'encargos de capacidade'. Eles têm sido adotados em alguns países como o Reino Unido, até a implantação do New Electricity Trading Arrangements (NETA) em março de 2001, a partir de quando a metodologia VLL/LOLP não mais foi aplicada, Argentina, Chile, Espanha e Colômbia. Quase sempre, tais pagamentos são feitos a cada gerador, com base na sua disponibilidade, tenha sido ele despachado ou não, ou com base na energia por ele gerada. O valor total a ser pago aos geradores é rateado por kWh da energia consumida, que pode ser colocado como um aditivo, aumentando o preço de mercado (*uplift*), como no caso do 'modelo inglês'.

2.3.1.1 Modelo Inglês

É um procedimento para o cálculo das taxas ou encargos de capacidade que se baseia nos estudos e índices de confiabilidade de longo prazo do sistema, principalmente a probabilidade de perda de carga, ou de interrupção do fornecimento, LOLP (*loss of load probability*), e em estimativas do VLL. Este processo é conhecido também como metodologia VLL / LOLP.

Neste modelo, a remuneração média de uma unidade geradora disponível e despachada, pela energia, em uma certa hora, é dada pela soma:

$$REM = CMS \cdot (1 - LOLP) + LOLP \cdot VLL \quad (2.1)$$

ou:

$$REM = CMS + LOLP \cdot (VLL - CMS) \quad (2.2)$$

onde:

CMS - custo marginal de operação de curto prazo do sistema em uma certa hora, em \$/MWh;

VLL - custo unitário da perda de carga, em \$/MWh (custo de interrupção).

A equação (2.1) mostra que a remuneração seria feita pelo custo marginal de operação com probabilidade igual a $(1 - LOLP)$ e pelo custo de interrupção com probabilidade igual à $LOLP$. Assim, caso a $LOLP$ atinja seu valor máximo 1, isto é, a perda de carga seja certa, a remuneração se torna igual ao VLL , como era de se esperar, mantendo a coerência com a proposta de mercado em que o custo marginal se iguala ao VLL em ocasiões de corte de carga, como visto na seção 2.2. Assim:

$$REM = VLL \quad (2.3)$$

O segundo termo de (2.2) representa o encargo, pagamento ou tarifa de capacidade unitária, pois representa um aumento do custo marginal de operação do sistema, caso a confiabilidade se encontre reduzida, ou seja, se houver maior probabilidade de não se conseguir atender a parte da carga. A expressão do pagamento pela capacidade despachada é portanto:

$$PC = LOLP \cdot (VLL - CMS) \quad (2.4)$$

No caso de usinas não despachadas, a remuneração seria composta apenas do pagamento por capacidade pois, como neste caso não haveria pagamento por venda de energia, (2.1) passaria a ser:

$$REM^{(ND)} = LOLP \cdot VLL \quad (2.5)$$

Em sistemas onde a energia é negociada em mercados abertos de curto prazo, o custo marginal do sistema é representado pelo preço de liquidação do mercado ou pelo preço de oferta de cada participante (*pay-as-bid*).

2.3.1.2 Pagamentos por Capacidade na Argentina

Na Argentina [CAMMESA, 2003] a capacidade é paga em certas horas do dia: das 9:00 às 24:00 h dos dias úteis; das 20:00 à 1:00 h aos sábados e das 20:00 às 23:00 h aos domingos e feriados. É definido um 'preço de mercado da potência', ou 'preço da potência posta à disposição' (\$/MWh), que é determinado em função de uma 'remuneração base unitária da potência', multiplicada por uma constante igual ou maior que a unidade, conforme (2.6):

$$\$PPAD = k_{PPAD} \cdot \$BASE \quad (2.6)$$

O valor da constante k_{PPAD} é definido pela Secretaria de Energia, em função da maior ou menor necessidade de se incentivar a implantação de nova geração.

A remuneração total por potência ou capacidade de cada gerador designado é dividida em duas parcelas:

- remuneração base da potência, calculada pelo valor unitário de (2.6);
- remuneração da reserva de curto e médio prazo, incluindo as reservas de quatro horas e de confiabilidade.

São remunerados apenas os geradores que tiverem sido designados através de um dispositivo denominado 'simulação para a remuneração base de potência' ou dos programas de despacho de reservas. Assim, na operação real do sistema, é possível haver máquinas suprindo potência, porém não previstas nesta simulação ou nestes programas, não contando assim com previsão de remuneração por tais serviços.

Pelo lado da demanda, na qual incluem-se os distribuidores, grandes consumidores, auto-produtores e contratos de fornecimento, incluindo os contratos de exportação, os encargos a serem pagos pela potência são de três tipos:

- pagamento pela potência despachada;
- pagamento pela reserva de potência;
- pagamento pelos 'serviços associados à potência'.

No caso da potência despachada, o pagamento é calculado em função da demanda do agente durante as horas em que a potência é remunerada. Para a reserva de potência, este pagamento é determinado em função de sua demanda máxima mensal. Quanto aos serviços associados à potência, o pagamento é calculado em função de sua 'potência declarada' e sua demanda máxima mensal.

Um gerador eventualmente pagará encargo por potência despachada, composto por duas parcelas, sendo a primeira formada pelos encargos que deveriam ser assumidos pela carga nos contratos de abastecimento nos quais ele vende energia, com esta transferência de responsabilidade, estabelecida nos contratos. A segunda é representada pela demanda correspondente à sua compra no mercado a vista (*spot*), para os contratos de exportação.

Ele deverá pagar também encargo por reserva de potência e/ou encargo por serviços associados à potência, caso possua contratos de abastecimento de venda de energia, nos quais assume o compromisso de se responsabilizar por estes encargos, em lugar da carga.

No que diz respeito aos serviços associados à potência, eles são constituídos pelas modalidades de reserva a seguir, definidas e comentadas no Apêndice B.

- reserva instantânea;
- reserva para regulação de frequência;
- reserva operativa de 5 minutos;
- reserva de 10 minutos;
- reserva fria de 20 minutos;
- reserva térmica de 4 horas;
- reserva de confiabilidade (reserva de médio prazo).

Excetuando-se o serviço de regulação de frequência, todos os demais são remunerados de forma adicional e independente da remuneração base de potência correspondente. A remuneração das reservas: operativa de 5 minutos, de 10 minutos e fria de 20 minutos é feita através de preços determinados com base no valor de \$PPAD, multiplicado por uma constante específica para cada uma delas. A reserva de confiabilidade, de forma distinta das reservas anteriormente citadas, que são designadas para as unidades pelo operador do sistema, é definida por licitação, sendo aceitas apenas as ofertas cujos preços são menores que um valor teto, calculado também com base no valor de \$PPAD, multiplicado por uma outra constante.

É importante observar-se então, que na Argentina existem remunerações específicas e independentes pela capacidade (potência) e por todas as modalidades de reservas, para regulação, operativas e de confiabilidade, sendo os preços de várias delas calculados em função do preço da potência.

2.3.1.3 Análise Custo x Benefício

Uma outra metodologia para a determinação das taxas de capacidade baseia-se numa avaliação da sua relação custo/benefício. O total a ser pago a título de encargo de capacidade, em um certo período de tempo, deve ser tal que o custo marginal de se incentivar maior confiabilidade seja igual ao benefício marginal que o sistema obtém deste aumento de confiabilidade. Apenas um certo número de unidades geradoras, selecionadas pelo critério de confiabilidade seria pago pela capacidade.

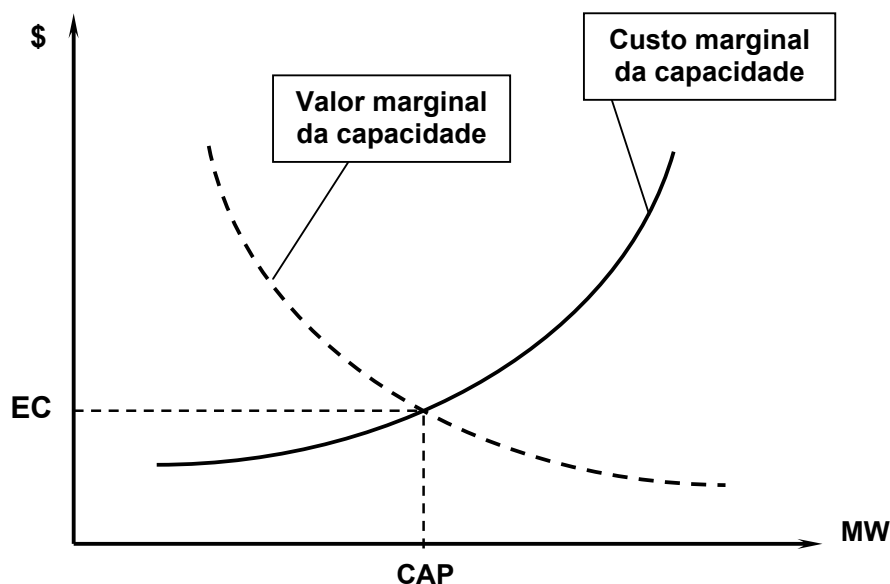


Figura 2.1 - Determinação do Encargo de Capacidade e da Capacidade Requerida

Desenvolveu-se um estudo comparativo [CEPEL, 1999] relacionado aos encargos de capacidade, o qual procedeu a uma avaliação dos valores do encargo de capacidade no sistema elétrico brasileiro para os anos de 2000, 2001 e 2002, utilizando a metodologia do modelo inglês e uma análise custo x benefício. Os valores unitários dos encargos de capacidade obtidos são mostrados na Tabela 2.1

Tabela 2.1 - Encargos de Capacidade Determinados pelo Estudo do Centro de Pesquisa de Energia Elétrica

	VALOR MÉDIO DO ENCARGO DE CAPACIDADE (US\$/MW)		
	2000	2001	2002
MODELO INGLÊS	4,83	2,19	0,86
ANÁLISE CUSTO x BENEFÍCIO	3,48	3,42	3,11

Neste estudo, como no caso da metodologia do modelo inglês, os encargos da demanda referentes à capacidade seriam calculados como:

$$PC_n = LOLP \cdot (VLL - CMS) \cdot L_n \quad (2.7)$$

sendo L_n a demanda de ponta da carga n (mensal ou anual).

Sugere-se que os encargos totais recolhidos, dados por:

$$PC_{TOT} = \sum_n PC_n \quad (2.8)$$

sejam rateados entre os geradores do sistema segundo a sua capacidade máxima de geração.

Para esta metodologia, não foi determinado o montante anual da capacidade a ser remunerada, por não haver uma definição de quais segmentos da carga arcariam com o seu pagamento. Em relação ao estudo de custo x benefício, este montante, definido como CAP na Figura 2.1, é encontrado de maneira automática, sendo seus valores aproximados mostrados na Tabela 2.2, juntamente com os demais resultados produzidos.

Tabela 2.2 – Demais Resultados do Estudo - Metodologia Custo x Benefício

ANO	Capacidade média a ser remunerada (MW)	Capacidade total a ser remunerada (GWh)	Encargo de capacidade total (US\$1000)	Impacto sobre o consumo (US\$/MWh)
2000	3480	10920	38000	0,10
2001	3420	9940	34000	0,09
2002	3110	6110	19000	0,05

O impacto sobre o consumo foi determinado considerando-se que ‘toda a energia demandada no Mercado Atacadista de Energia (MAE)’ seria responsável pelo pagamento dos encargos, o que não esclarece bem quais consumidores estariam sujeitos a eles.

2.3.2 Tarifas com Componente de Capacidade

As tarifas de dois componentes ou ‘binômias’, foram propostas pela primeira vez em 1892 por John Hopkinson, originalmente como *maximum-demand tariffs*, planejadas com base na consideração de que os custos fixos que poderiam ser atribuídos a um certo consumidor são proporcionais à sua carga máxima (kW) e os custos variáveis ao seu consumo (kWh).

Em termos de composição, as atuais tarifas binômias permanecem com a característica original, compostas de valores relativos ao consumo (\$/kWh) e à demanda (\$/kW). Elas podem incorporar ainda valores diferenciados para a demanda, relacionados ao nível de demanda do sistema e ao nível de precipitação pluviométrica, ou estação do ano. São definidos limites para a demanda, a partir dos quais a tarifa específica sofre um considerável aumento.

2.4 Experiência Mundial e os Problemas Verificados

Ao se adotar no Reino Unido os pagamentos específicos pela capacidade, verificou-se que, não obstante em ambientes perfeitamente competitivos existirem argumentos bem fundamentados para a elevação dos preços da energia, através da adição de parcelas relativas ao pagamento da capacidade, observaram-se abusos nestes últimos, causados em geral pelo poder de mercado excessivo.

Em primeiro lugar a metodologia tradicional do cálculo da LOLP, exagerando a probabilidade de usinas não estarem disponíveis, levando assim ao aumento dos pagamentos por capacidade [OFGEM, 1999]. Em função disto, geradores passaram a fornecer declarações inexatas de sua indisponibilidade, recebendo pagamentos pela capacidade baseados em previsões incorretas da capacidade ociosa do sistema. Mesmo com a mudança da metodologia de cálculo da LOLP, alguns geradores, em certas circunstâncias, passaram a atrasar as revisões de declaração de indisponibilidade que devem ser feitas após um gerador experimentar uma falta, com o intuito de ampliar os valores recebidos por capacidade.

As principais conseqüências decorrentes foram o aumento irreal do valor unitário do pagamento pela capacidade e o ganho de vantagens individuais indevidas por alguns geradores. Para que se tenha uma idéia da possibilidade de especulação visando influenciar os pagamentos por capacidade, basta dizer que no ano fiscal de 1994/1995, estes pagamentos representaram 20% do pagamento total por energia gerada, o que daria para reconstruir 6% de toda a capacidade instalada no Reino Unido.

Segundo o Office of Gas and Electricity Markets [OFGEM, 1999]:

“Pagamentos por capacidade são adotados com o objetivo de fornecer sinais relativos à necessidade de aumento de capacidade, tanto em curto-prazo quanto em longo prazo. Entretanto, eles não fornecem um sinal de curto-prazo muito efetivo para encorajar a geração e a demanda a responder às circunstâncias, variando rapidamente, uma vez que eles não refletem exatamente as variações de curto prazo na margem de capacidade. Além disso, de ano para ano, os pagamentos por capacidade não funcionaram como pretendido: em anos em que os pagamentos por capacidade foram baixos, os geradores aumentaram os preços marginais do sistema. Em adição, estes pagamentos complexos e administrados fornecem um sinal de longo prazo muito fraco para a necessidade de capacidade”.

Na Califórnia não se estabeleceram pagamentos por capacidade, adotando-se um mercado apenas de energia, sem componente de capacidade, além de mercados de serviços ancilares. Os pagamentos com alguma afinidade com aqueles feitos pela capacidade, são concedidos apenas aos geradores ofertantes de reservas, vendidos nos mercados específicos, o que mostra a opção pela política de *operating-reserve pricing*. Estes pagamentos são definidos através de leilões competitivos, o que teoricamente tornaria mais difícil a especulação. Entretanto, os mercados nem sempre funcionaram competitivamente, conforme era de se esperar. Os preços nos mercado mostraram-se altamente voláteis, mesmo em períodos longos de demanda invariável. Os preços de reservas de qualidade inferior (substituição) superaram os preços de reservas de maior qualidade (regulação) e muitas vezes os preços das reservas chegaram a superar os preços da energia. Foram ainda observados eventos de especulação nas ofertas de capacidade dos geradores.

No sistema argentino, generosos pagamentos por capacidade, determinados com base na energia gerada, levaram algumas empresas geradoras a fazer ofertas de energia abaixo do custo marginal, com o claro objetivo de expandir a energia gerada, e assim aumentar a receita com os pagamentos por capacidade.

2.5 Encargos de Capacidade no Brasil

2.5.1 Propostas do Modelo de Reestruturação do Setor Elétrico

No processo de reestruturação do setor elétrico no Brasil, iniciado em 1998, o pagamento por capacidade foi abordado em diversas oportunidades conforme relatado em seguida.

Segundo a versão inicial das Regras do MAE [MAE,1998]:

“O preço do MAE deverá incorporar um encargo por capacidade, com o objetivo de recuperar os custos fixos de manutenção e operação dos geradores existentes, que tenham sido solicitados para manter um nível de confiabilidade adequado no sistema. O encargo por capacidade deverá ser pago a todos os geradores que se declararem disponíveis para o dia seguinte. Este encargo deverá ser estabelecido *ex-ante* e projetado para incentivar os geradores a se tornarem disponíveis quando forem mais necessários à operação do sistema.”

Ainda sobre o encargo de capacidade encontra-se neste mesmo documento [MAE,1998]:

“Um encargo por capacidade mínimo, expresso em R\$/MW será pago aos geradores como parte do preço do MAE. Deverá ser o mais simples possível, e destinado a minimizar as distorções aos incentivos. Particularmente, o encargo por capacidade deverá:

- cobrir apenas os custos fixos de operação e de manutenção de uma planta, assumindo baixa total de seus ativos. O encargo por capacidade não será utilizado para gerar sinais para investimentos em nova capacidade;
- remunerar os geradores por terem estado disponíveis quando esta disponibilidade era de grande importância para o sistema;
- incentivar as decisões apropriadas de desativação das plantas”.

Na ocasião da publicação destas regras não se havia chegado ainda a uma forma precisa para a implementação deste mecanismo. No entanto, tem-se ainda [MAE,1998]:

- “será provavelmente um valor variável conforme a estação do ano e deverá estar concentrado nos períodos de ponta de cada dia, para incentivar a disponibilidade quando o sistema estiver com maior necessidade de capacidade. A variação horosazonal deverá ser determinada antecipadamente pelo Operador Nacional do Sistema (ONS), utilizando um modelo apropriado do sistema e deverá ser aprovada pela Agência Nacional de Energia Elétrica (ANEEL), para vigorar pelo período de um ano;
- o encargo por capacidade será pago somente aos geradores que se declararem disponíveis no dia anterior ao despacho. Geradores que não se declararem disponíveis neste estágio, mas que o fizerem através de uma redeclaração posterior, não farão jus ao recebimento deste encargo;
- o nível do encargo por capacidade poderá ser revisto anualmente para permitir aumentos nos pagamentos, caso seja previsto um ano seco;
- sua implementação deverá ser acompanhada de uma análise para assegurar que o valor da carga não atendida seja estabelecido em bases compatíveis com o encargo por capacidade, isto é, que o efeito conjunto do pagamento pela capacidade com o valor da carga não atendida, proporcione o nível desejado de pagamento aos geradores e de confiabilidade do sistema;
- deverá ser revisto antes do término do 4º ano do início da operação do MAE. Esta revisão irá permitir abordagens alternativas que deverão ser consideradas antes que os contratos iniciais sejam extintos”.

O assunto foi ainda discutido publicamente, em um processo de audiência pública [ANEEL, 2000a] não se chegando, entretanto, a qualquer consenso ou decisão sobre a necessidade ou adequação da implementação deste pagamento, mantendo-se ainda dúvidas a respeito da sua eficiência em atingir os objetivos com ele pretendidos.

Na versão II das Regras Algébricas do MAE, são encontradas ainda abordagens do encargo de capacidade no Capítulo 11, “Cálculo do Encargo de Capacidade” [MAE, 2000a] e no seu texto descritivo e ainda, no apêndice G, “Especificação da Modelagem do Encargo de Capacidade” [MAE, 2000b], onde é proposta uma metodologia específica

para o cálculo deste encargo. Além disso, o assunto é tratado também nos Capítulos 3 e 9 e os seus respectivos textos descritivos.

Já pela versão final das Regras da 1ª Etapa do MAE [MAE, 2001], implantada pela Resolução ANEEL nº 290/2000 [ANEEL, 2000b] também no Capítulo 11: “Cálculo do Encargo de Capacidade”, conclui-se que a decisão à respeito deste encargo seria postergada, uma vez que neste capítulo e em seu texto explicativo encontram-se respectivamente, apenas as afirmações: “O Capítulo de Encargo de Capacidade não será utilizado nestas Regras de Implementação do MAE” e “Este encargo não será utilizado nas Regras de Implementação do MAE”.

Na resolução ANEEL nº 290/2000 [ANEEL, 2000b], fica claro entretanto, que na 3ª etapa da implantação das regras do MAE, deveria ser possibilitada a implantação dos encargos de capacidade, conforme visto no seu Artigo 13:

“A forma de cálculo do encargo de capacidade, objeto do Capítulo 11 das Regras do MAE deverá refletir as premissas constantes do “Documento Básico para o Estabelecimento das Regras do Mercado”, devendo ser submetida pelo MAE à homologação da ANEEL visando possibilitar a sua implantação a partir da 3ª etapa”.

No entanto, nas versões posteriores das Regras Algebricas de Mercado do MAE [MAE, 2002a; MAE, 2002b; MAE, 2003a; MAE, 2003b; MAE, 2003c; MAE, 2003d] não mais se prevê a implantação dos encargos de capacidade.

Sem os encargos de capacidade, o que teoricamente ocorreria é que os geradores que operam com altos custos variáveis receberiam uma remuneração extremamente alta, mas apenas durante um evento de interrupção de fornecimento, o que compensaria todo o período sem remuneração, pelo menos em teoria. Como este tipo de evento tem probabilidade de ocorrência muito baixa a remuneração dos geradores 'de ponta' se torna incerta, criando-se ainda uma alta volatilidade nos seus pagamentos. Esta volatilidade poderia, além de incentivar desativações, desencorajar prováveis investidores, comprometendo a expansão do sistema. O principal objetivo de se implantar os pagamentos destes encargos pela carga, seria a diminuição desta volatilidade e a garantia de uma receita mais regular para as empresas geradoras, que as permitam manter as suas usinas de ponta em operação. No Brasil foi apresentada inicialmente uma metodologia [MAE, 2000b] para o cálculo destes encargos, seguida de uma segunda, por

parte do Comitê de Revitalização do Modelo do Setor Elétrico (CRMSE), que possui muito pouca diferença em relação à primeira, conforme apresentado no Apêndice C

2.5.2 Proposta do Comitê de Revitalização do Modelo do Setor Elétrico

Mais recentemente, no conjunto de medidas propostas para a revitalização do modelo do setor elétrico, é feita uma nova abordagem ao encargo de capacidade. No item referente às Regras do MAE é feita uma proposta de revisão completa destas regras, envolvendo a solução das questões ainda hoje pendentes, no que se inclui o próprio encargo de capacidade, considerado uma regra anteriormente rejeitada.

O assunto pode ser encontrado em vários pontos do Relatório de Progresso nº 2 [CRMSE, 2002], do Comitê de Revitalização do Modelo do Setor Elétrico, por exemplo nos itens: Estímulo à Contratação de Reserva de Geração e Expansão da Capacidade de Ponta, e mais detalhadamente, no documento de apoio I [CRMSE, 2002a]. Em resumo, é novamente feita uma recomendação no sentido de se introduzir a cobrança de encargo de capacidade no Brasil, mantendo-se a orientação da Resolução ANEEL nº 290/2000 [ANEEL, 2000b].

É importante observar-se que a conotação dada à reserva de geração pela proposta de revitalização, é de uma reserva de longo prazo, através da contratação de reserva de geração e da expansão da capacidade de ponta do sistema, visando solucionar mais as questões de racionamento de energia a longo prazo e o possível aumento da demanda reprimida no horário de ponta. Ela deve ser entendida portanto, como uma reserva de planejamento.

Algumas outras observações relacionadas aos encargos de capacidade, abordadas no documento de apoio I [CRMSE, 2002a], segundo a metodologia proposta, são colocadas a seguir:

- eles não aumentam os custos médios do sistema, mas simplesmente reorganizam o fluxo financeiro dos pagamentos, evitando a sua concentração em alguns períodos e distribuindo-os ao longo do tempo;

- o valor que cada gerador recebe a título de encargo de capacidade é igual ao valor esperado da sua remuneração no MAE, ao se simularem todos os eventos capazes de levar à uma interrupção de fornecimento;
- o montante anual total a ser recebido pelos geradores, seria pago por toda a carga, de maneira uniforme, ao ser somado aos encargos de serviços do sistema. Nas simulações realizadas, este valor seria de aproximadamente um bilhão de dólares para todo o país.
- uma objeção em relação a este processo de pagamento é que ele impede geradores e consumidores de se protegerem contra a sua própria volatilidade, através dos contratos bilaterais, na medida em que é feito fora do preço do MAE;
- as empresas distribuidoras e comercializadoras que assinaram os contratos iniciais estariam, em princípio, sujeitas ao pagamento deste encargo, apesar dos textos dos mesmos estabelecerem claramente que os pagamentos contratados já englobavam um pagamento pela capacidade, além da energia a ser fornecida;
- a isenção destes contratos do encargo de capacidade não resolveria plenamente esta questão, pois ela persistiria em função da natureza dos contratos novos, que seriam assinados em um ambiente de livre negociação. Seria importante que estes contratos se mantivessem como um mecanismo de proteção contra a volatilidade futura dos encargos relativos à capacidade.

Uma solução para as questões suscitadas anteriormente seria a incorporação de todo o encargo de capacidade, ou pelo menos a maior parte possível dele, no preço do MAE. Um procedimento análogo é adotado no modelo de reestruturação, inicialmente implantado no Reino Unido, conforme visto na Seção 2.3.1.1.

A metodologia proposta pelo CMRSE é mostrada com detalhes no Apêndice C.

2.6 Reservas de Planejamento e Reservas Operativas

2.6.1 Introdução

As reservas de planejamento, ou 'reservas de capacidade instalada', têm a finalidade de contribuir para que o risco de não se ter recursos suficientes para satisfazer as demandas dos clientes seja mantido em um nível aceitável. Em outras palavras, estas reservas são planejadas para assegurar que capacidade suficiente esteja disponível para atender as cargas dos consumidores⁴, em um horizonte de longo prazo, levando-se em conta:

- saídas planejadas para manutenção;
- saídas forçadas prolongadas;
- 'aposentadorias' inesperadas de unidades, devido à falhas em equipamentos;
- saídas de unidade para modificação ou aumento de capacidade;
- demandas inesperadas dos consumidores;
- saídas de transmissão que influenciem o fornecimento de energia;
- condições climáticas incomuns;
- atos da natureza.

Quanto às reservas operativas, essas estão ligadas à preservação da segurança do sistema elétrico, conforme discutido no Apêndice A. Elas são previstas com a finalidade de manter a integridade do sistema em um horizonte de curto prazo, ou seja, no dia-a-dia da operação e serem mantidas por períodos curtos, de até 4 horas. Para isto elas devem garantir a continuidade da entrega da energia necessária, em seguida a perdas de geração, de importações ou de recursos de transmissão. Reservas operativas permitem aos operadores do sistema administrar perturbações de curto prazo e re-equilibrar rapidamente o sistema em seguida a uma contingência. Essencialmente, o propósito das reservas operacionais é fornecer resposta imediata e em curto prazo a situações de contingência operacional. Dentro das reservas operacionais incluem-se também as reservas para regulação de frequência que, na ocorrência de uma contingência, apesar de terem um objetivo específico, são as primeiras a serem utilizadas, devendo ser repostas em seguida.

⁴ Existe ainda uma outra necessidade de reservas, relacionada à energia, que tem a ver com os níveis dos reservatórios das usinas hidráulicas e com a disponibilidade de combustíveis para as usinas térmicas. A sua consideração e análise não será objeto do presente trabalho.

As reservas de planejamento e as reservas operativas relacionam-se portanto, respectivamente, com os atributos de adequação e segurança, componentes da confiabilidade dos sistemas elétricos de energia [Oren, 2000; NPPC, 2001].

2.6.2 Interação entre as Reservas de Planejamento e as Reservas Operativas

Analisando pelo aspecto técnico, é fácil concluir-se que a adequação e a segurança possuem um relacionamento próximo, uma vez que sistemas com abundância de reserva de capacidade (planejamento), proporcionam maior flexibilidade no tratamento de perturbações imprevistas [Oren, 2000], sem entretanto, garantir a segurança. Além disso, elas não são excludentes, já que um único gerador pode satisfazer a ambos os requisitos de maneira complementar e, eventualmente, as reservas de planejamento podem ser utilizadas em reposição às reservas operacionais. No entanto, algumas considerações exigem que elas sejam vistas de forma diferente.

O que se observa em um sistema elétrico em um certo momento, é uma combinação de unidades geradoras despachadas, não despachadas ou parcialmente despachadas, o que determina um montante de geração ociosa distribuído entre algumas usinas, que poderia ou não ser tratada como reserva, de um ou outro tipo. Existem vários motivos para uma unidade geradora não ter sido despachada, total ou parcialmente:

- manutenção preventiva programada;
- manutenção corretiva não programada;
- custos de operação mais altos;
- restrições causadas por problemas hidrológicos;
- restrições causadas por problemas elétricos;
- restrições causadas por problemas ambientais.

Isto, somado à possibilidade da presença de diversas tecnologias de geração, permite que ocorra uma infinidade de condições operacionais, que fazem com que os tempos para disponibilização para uso das diversas parcelas desta geração não despachada, sejam muito variáveis. O inter-relacionamento entre os vários segmentos das reservas operativas e as reservas de planejamento e a sua superposição são mostrados na Figura 2.2, na qual este tempo cresce de baixo para cima, na faixa da disponibilidade não despachada.

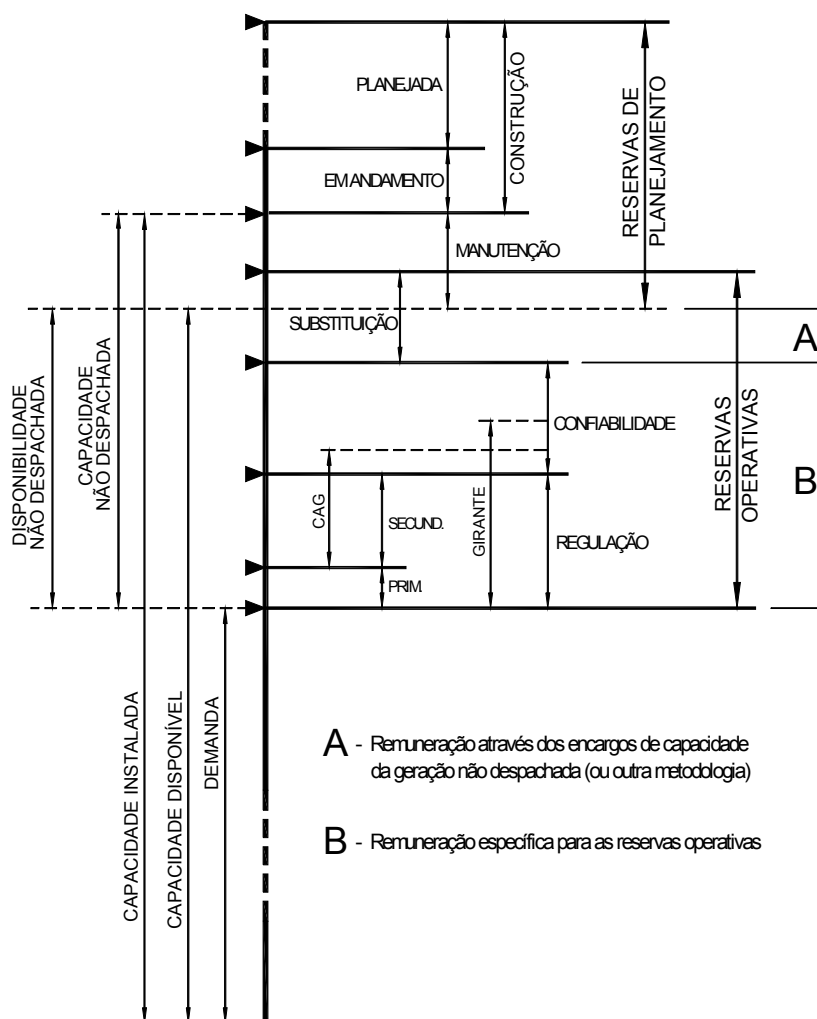


Figura 2.2 - Reservas de Planejamento e Operativas - Relacionamento e Remuneração

Reside nisto o principal fator para se conseguir a diferenciação entre as reservas que poderiam ser consideradas como operativas e aquelas que não. Como as reservas operativas exigem menores tempos para disponibilidade de uso, unidades em manutenção, ou com restrições operacionais de solução demorada, não se prestariam a fornecer este tipo de reserva. Entretanto, a sua geração ociosa poderia ser caracterizada como uma reserva de planejamento, de longo prazo. O mesmo pode ser dito para usinas em fase de testes ou de construção, que comporiam também as reservas de planejamento, dependendo do prazo para a sua colocação em operação.

Este aspecto relacionado ao tempo é uma condição necessária, mas a rapidez com que se conseguiria transformar a reserva em energia não é suficiente para definir se uma

certa parcela de geração ociosa será considerada como reserva de potência operativa. Existem outros tipos de problemas que podem impedir uma unidade geradora de fornecer reservas operativas. Podem ser citados, por exemplo, as mesmas restrições operacionais e de custos, listadas anteriormente. Além do mais, certas modalidades de reserva operativa exigem o concurso de equipamentos e instalações apropriadas, como as reservas de regulação.

Desta forma, pode-se afirmar que uma unidade geradora somente se tornaria fornecedora de reservas operativas, após ser designada para tal pelo operador do sistema, e a partir do momento em que se colocasse operacionalmente condicionada e pronta para fazê-lo. Esta 'prontidão' exige que a usina ou unidade geradora tenha um tratamento diferenciado, com uma monitoração mais cuidadosa, sob o ponto de vista da operação e manutenção. A segurança do sistema estaria dependendo parcialmente desta condição, e a usina ou unidade estaria sendo vista pelo operador de uma forma especial.

Todas estas colocações contribuem para justificar a caracterização das reservas operativas como um serviço completamente independente e diferenciado das reservas de longo prazo, sob uma visão técnica. Esta diferenciação é reforçada, também pela definição dos índices de confiabilidade, como a LOLP. Além da diferença temporal, que influencia no valor dos índices, no longo prazo este índice pode considerar informações relacionadas às séries hidrológicas e níveis de reservatórios, na avaliação da capacidade de geração do sistema. No curto prazo geralmente só se consideram informações sobre as taxas de falha das unidades geradoras e elementos de transmissão e a carga prevista.

2.6.2.1 Análise do Aspecto Econômico Envolvido

O pagamento por capacidade pode ser considerado como o pagamento pelas reservas de planejamento, ou pela capacidade disponível, de forma indireta, na medida em que pode ser encarado basicamente como um incentivo a que haja uma expansão da capacidade de geração do sistema, ou seja, a que as unidades existentes aumentem a sua disponibilidade de geração e que novas unidades geradoras se estabeleçam. Pelo menos este é um dos resultados esperados pelos defensores de seu uso. Isto permite que tal pagamento, juntamente com o pagamento pelas reservas de potência operativa, sejam tratados como 'pagamentos pela confiabilidade do sistema elétrico'.

Assim, paralelamente a sua interação sob o aspecto técnico, a relação de ambas as modalidades de reservas, de planejamento e operativa, com o quesito confiabilidade, poderia vir a fazer com que elas se confundissem, também sob o aspecto econômico, com conseqüências talvez mais graves. Algumas questões econômicas fundamentais surgiriam:

- poderiam os pagamentos por capacidade (encargos de capacidade) serem únicos, considerando-se que eles já estariam remunerando também as reservas de potência operativa?
- caso se pague uma unidade geradora por reserva de potência operativa, não estaria havendo uma duplicidade de pagamento, uma vez que esta unidade poderia já ter recebido um pagamento por toda ou por parte da sua capacidade disponibilizada, advindo do encargo de capacidade, ou seja, um pagamento como reserva de planejamento?
- caso esta duplicidade ocorra, seria ela justa, isto é, ela deveria ser aceita, considerando tratar-se de dois serviços diferentes?
- se aceita para alguns agentes, ela não poderia suscitar questionamentos de agentes que estivessem sendo pagos pela prestação de um ou de outro serviço, no sentido de também terem o direito de receber pelos dois?
- se aceita, ela não poderia incentivar a utilização de procedimentos pouco recomendáveis, por parte de certos agentes, para garantir receitas maiores, conforme já se observou?

É importante portanto, que se analise com cuidado esta possível interferência entre estas reservas, para que possam ter caracterização e quantificação adequadas, além de um gerenciamento eficiente, de maneira a serem correta e justamente remuneradas pelo benefício que cada uma proporciona ao sistema elétrico. Para isto é necessária a investigação da superposição das faixas de reservas de planejamento e de reservas operativas, nas suas diferentes modalidades, com o objetivo de se definirem as fronteiras entre elas e o alcance de cada uma, o que pode ser visto na Figura 2.2.

As reservas de planejamento proporcionam adequação e envolvem questões de longo prazo relacionadas ao planejamento da geração e investimentos, sendo o custo marginal a elas relacionado, o 'custo marginal de longo prazo', ou seja, o custo de se aumentar a capacidade de geração em 1 MW. As reservas operativas, por seu turno, proporcionam segurança e tratam com questões de curto prazo, relativas à operação de despacho de geração [NPPC, 2001], sendo o seu custo marginal o 'custo marginal de curto prazo', ou seja o custo de tornar disponível 1 MW de potência (reserva).

2.7 Tarifas de Energia Elétrica e Custos Fixos

Um dos principais componentes dos custos fixos relacionados à indústria de energia elétrica é representado pelo custo do investimento de capital. O setor de geração de energia elétrica caracteriza-se pela utilização intensiva de capital, que pode chegar a 400% da receita anual [Bitu, 1993], o que faz com que a fixação das taxas de retorno seja de extrema importância, principalmente para os produtores de energia.

Neste sentido, a definição das tarifas da energia elétrica vem a exercer um papel crucial sobre as condições financeiras das empresas distribuidoras e comercializadoras, refletindo-se em consequência, sobre a saúde financeira das empresas de geração e transmissão. É natural que a convivência por muito tempo com tarifas de energia elétrica abaixo dos custos totais de produção provoque efeitos devastadores sobre as finanças destas empresas.

No enfoque tradicional, a definição de tarifas no setor elétrico tem sido feita com base no custo do serviço prestado, a partir dos custos contábeis, procedimento também denominado de 'regulação da taxa interna de retorno'. Neste caso, garante-se a recuperação dos custos de geração, operação e manutenção, dos custos de depreciação dos equipamentos e instalações, além de uma rentabilidade satisfatória sobre o capital investido. Assim, um aspecto muito importante deste enfoque, é a garantia da recuperação dos custos fixos, principalmente aqueles relacionados aos investimentos de capital. A regulação da taxa de retorno é feita para se evitem lucros excessivos, funcionando na verdade como uma forma indireta de determinação de preços [Pires, 1994].

O componente de destaque da tarifa, neste enfoque, seriam os custos de capital, calculados diretamente em função do capital imobilizado e da rentabilidade. Para se reforçar a importância da questão dos custos fixos e da sua remuneração, deve ser considerada a necessidade de se definir corretamente o valor do capital imobilizado das empresas, ou seja, o investimento sobre o qual se aplicaria a taxa de retorno (*sunk costs*) e a grande dificuldade em fazê-lo. Sob a ótica da economia, o custo de oportunidade do capital deve determinar a taxa de retorno. A definição destes valores com base em custos históricos veio a trazer grandes prejuízos para algumas empresas, em períodos de maior inflação, devido à desvalorização dos seus ativos [Breyer, 1982].

A tarifação através dos custos dos serviços apesar de ser facilmente adaptável aos objetivos do fornecimento de energia elétrica como serviço público e evitar abusos de preços, possui o grave defeito de não garantir uma eficiência produtiva pelo próprio fato de garantir a cobertura de todos os custos, assegurando antecipadamente uma razoável taxa interna de retorno. Isto leva a uma deterioração da qualidade do serviço, além de desincentivar os investimentos em nova geração.

Um mecanismo que reconhecidamente possui capacidade de levar a um emprego mais eficiente dos recursos financeiros, conseguindo maior eficiência econômica, é a determinação de preços da energia e a tarifação do consumo com base nos princípios dos custos marginais. O uso de custos marginais, apesar de algumas desvantagens e dificuldades de implementação, possui muitas vantagens, como o aumento tanto da eficiência alocativa quanto do aproveitamento da capacidade instalada devido a um melhor gerenciamento da demanda.

O princípio básico da tarifação pelo custo marginal é a alocação ao consumidor dos custos adicionais em que o sistema incorre ao atendê-lo. Como cada categoria de consumidores impõe custos diferentes, as tarifas são diferenciadas primeiramente em função destas categorias, ou seja, conforme o objetivo do uso da energia, se industrial, comercial, rural, residencial etc. Outros fatores diferenciadores são também considerados, citando-se:

- níveis de tensão: alta, média e baixa tensão;
- horários da ocorrência do consumo: no período de ponta de demanda do sistema ou fora dele ;
- períodos do ano e condições climáticas: período seco e período úmido;

- regiões geográficas etc.

De posse destas informações, qualitativas e quantitativas do comportamento da demanda, é possível a identificação dos custos marginais do fornecimento, com a consideração inclusive da influência do aspecto probabilístico destes custos no comportamento do processo de oferta e de demanda de energia.

Entretanto, neste novo enfoque tarifário com base em custos marginais, proposto e adotado em função dos processos de reestruturação do setor elétrico, caracterizados principalmente pela mudança de regulamentação, observam-se também dificuldades relacionadas mais uma vez com a questão da cobertura dos custos fixos da indústria elétrica. Os grandes investimentos de capital, típicos do setor elétrico, exigem uma indispensável e correta distribuição dos custos incrementais entre os consumidores, principalmente os fixos, sob pena de se inviabilizarem tais investimentos. Os preços definidos com base apenas no custo marginal não promovem esta distribuição pois, por definição, os custos marginais independem dos custos fixos. Assim poderia haver prejuízos para as empresas, de forma que, para evitá-los é necessário definirem-se procedimentos específicos para se resolver como os custos fixos serão tratados.

Uma primeira proposta para solução desta questão seria através da cobrança de um valor adicional ou taxa extra, com o objetivo de se cobrirem os custos fixos. Esta modalidade de intervenção é criticada, pois pode introduzir aspectos ineficientes na estrutura tarifária, como a possibilidade de excluir consumidores de baixa renda do mercado, uma vez que as preferências dos consumidores são diferenciadas e desconhecidas [Armstrong, 1994].

Outra opção para a distribuição dos custos fixos seria a utilização da Regra de Ramsey (*Ramsey Pricing Rule*) que propõe que esta distribuição entre os diversos produtos ou postos tarifários seja feita por *mark-ups* sobre os custos marginais, na proporção inversa das elasticidades das demandas dos consumidores [Breyer, 1982; Viscusi, 1995].

Outras possíveis imperfeições associadas à perfeita aplicação de uma tarifação baseada na teoria dos custos marginais, seriam [Pires, 1994]:

- a assimetria das informações prestadas pelas empresas participantes do setor;

- o maior custo da energia, devido ao acréscimo dos custos fixos nos horários de ponta, com prejuízo para os consumidores com grande consumo nestes horários, e que não têm como alterar seus processos produtivos;
- os altos custos envolvidos nos estudos técnicos e de viabilidade econômica e na instalação de medidores digitais, necessários e adequados a esta metodologia;
- as dificuldades para a obtenção de modelos confiáveis de previsão de elasticidades e de curvas de demanda.

Devem ser lembradas ainda as questões regulatórias restritivas com as quais a tarifação pelos custos marginais se defronta, típicas dos serviços públicos. Entretanto, a despeito de quaisquer dificuldades, as estruturas tarifárias baseadas no custo marginal constituem uma evolução profunda, sob a ótica da eficiência econômica, quando comparadas com a tarifação baseada em custos do serviço. As principais contribuições deste novo modelo tarifário seriam um gerenciamento bem mais eficiente da demanda e uma sinalização para os preços da energia em segmentos desregulados da indústria de energia elétrica, como o mercado à vista, facilitando a abertura de espaço para a competitividade no setor [Pires, 1994].

Como se conclui, a grande discussão proporcionada pela adoção de um modelo tarifário mais eficiente reside também na necessidade de remuneração dos custos fixos, a qual exerce influência na capacidade de geração instalada dos sistemas elétricos, devido a sua relação direta com a potência ou capacidade disponível.

2.7.1 Tarifas de Fornecimento de Energia Elétrica na Indústria Brasileira de Eletricidade

No Brasil, até o ano de 1981, as tarifas de fornecimento podiam ser classificadas basicamente em duas categorias [Pires, 1994; Bitu, 1993]:

- tarifas que consideram apenas a energia consumida, ou tarifas monômias, e
- tarifas com dois componentes, de energia ou tarifa de consumo e de potência ou tarifa de demanda, denominadas tarifas binômias ou tarifas de Hopkinson.

A partir daquele ano, começaram a ser feitas considerações em relação aos aspectos de maior eficiência econômica na utilização dos custos marginais na definição da estrutura tarifária, através da criação das tarifas binômias conhecidas como 'horo-sazonais'. Esta denominação deve-se ao fato destas tarifas incorporarem as variações nos custos marginais, conforme já abordado, por influência principalmente da hora do dia e da estação do ano em que o consumo é feito.

Sob o aspecto de avaliação e quantificação dos valores finais a serem pagos pelos consumidores, são levantadas duas características fundamentais do consumo, além do suporte de potência reativa (KVAR):

- energia total consumida - kWh;
- demanda, ou taxa do fluxo de energia por unidade de tempo - kW (potência máxima requerida),

No seu componente de potência ou tarifa de demanda, a estrutura tarifária horo-sazonal define através de contrato, valores permitidos de demanda, tanto para os períodos de ponta quanto para os períodos fora de ponta. Além disso, tais valores são ainda diferenciados para períodos 'seco' e 'úmido' do ano. As maiores tarifas de demanda são observadas nos casos da chamada 'ponta seca' e a não observação dos valores contratuais de demanda implicam na aplicação de outros valores de tarifas, denominados 'tarifas de ultrapassagem', financeiramente muito mais pesadas para o consumidor.

Juntamente com a pretendida eficiência econômica, pelo lado dos agentes geradores e distribuidores de energia, pretende-se também com tal estrutura tarifária, nivelar a demanda de energia ao longo do tempo, aumentando o fator de carga global, que é a relação entre a demanda média e a demanda máxima verificadas no período, aproximando o seu valor da unidade. Com isso o custo unitário do fornecimento de um valor médio de energia poderia ser reduzido de forma significativa, conseguindo-se ainda a redução da razão entre o capital necessário e a receita prevista de um empreendimento elétrico [Bitu, 1993]. O fator determinante deste possível nivelamento, ou deslocamento da ponta de consumo, é exatamente a diferenciação de preços entre os períodos diários considerados de ponta ou não, complementada pela diferenciação das condições sazonais de precipitação pluviométrica.

Resumindo, as inovações mais visíveis foram, além das considerações horo-sazonais, a introdução da tarifa de demanda com influência distinta em relação aos quesitos operacionais dos sistemas de energia elétrica.

É bem verdade que, atualmente, as tarifas com componente de potência ou demanda não são aplicadas para a totalidade dos consumidores, ficando de fora, por exemplo os consumidores residenciais e consumidores comerciais e industriais de pequeno porte (baixa tensão), o que significa que nem toda a carga dos sistemas vem pagando às distribuidoras pelo componente de demanda, apesar destas pagarem este encargo às empresas geradoras. Isto se dá principalmente pela impossibilidade de se proceder em curto prazo à medição da demanda destes consumidores, por questões de logística e de custos conforme já mencionado. Apesar destas dificuldades, as empresas distribuidoras têm adotado procedimentos para implementar esta medição para todos as modalidades de consumidores, como a instalação de medidores eletrônicos digitais modernos nos novos consumidores comerciais e industriais. Com este mesmo objetivo encontra-se atualmente em fase de estudos, a chamada 'tarifa amarela', a qual viria a permitir o uso de tarifas binômias e horárias para consumidores de baixa tensão, mesmo sem a instalação de novos medidores.

2.7.2 Relacionamento entre Tarifas, Custos Fixos e Confiabilidade

A confiabilidade de um sistema elétrico é uma séria fonte de custos, explicada em parte por exigir no sistema, durante a maior parte do tempo, um excesso de capacidade instalada. Este excesso só estará presente caso haja razoável cobertura dos custos fixos relacionados ao capital investido, pois é sabido que a deficiência na remuneração dos custos fixos produz uma sinalização negativa aos investimentos na capacidade de geração, com prejuízo da confiabilidade global do sistema elétrico. Os preços e tarifas da energia elétrica, além de economicamente eficientes no sentido de proporcionarem uma rentabilidade satisfatória, devem ainda garantir uma confiabilidade adequada no atendimento da demanda de energia.

Em relação à tarifação pelo custo do serviço, apesar dela garantir a cobertura dos custos fixos, é possível que venha a trazer redução de confiabilidade, já que quando se tem custos crescentes e preços definidos a partir dos custos unitários, uma indústria monopolista tem tendência a limitar a sua produção em um nível no qual se tenha a

igualdade entre o custo marginal e o preço, o que pode resultar em produção inferior à demanda para este preço [Bitu, 1993].

Quanto às tarifas horo-sazonais, baseadas em custos marginais, deve ser destacada a clara preocupação com o quesito de confiabilidade, em um horizonte de médio a longo prazo, agregada pelo seu componente de demanda ou de capacidade. Em primeiro lugar, este componente proporciona sinalização econômica à expansão da capacidade de geração, e assim da confiabilidade, ao ser uma fonte de cobertura para os custos fixos de capital. Esta preocupação é reforçada pela definição de valores de tarifa maiores para as demandas em horários coincidentes com a ponta de demanda do sistema e pelas pesadas tarifas de ultrapassagem. O nivelamento da curva de carga pretendido pela distinção dos períodos diários de ponta, pode ser interpretado como um aumento das reservas de potência, principalmente nestes períodos, ou seja, uma forma adicional de se conseguir maior confiabilidade. As tarifas de ultrapassagem, por sua vez, tendem a tornar mais eficiente o uso da energia, na medida em que, além de obrigarem a uma programação otimizada do uso das cargas pelos consumidores, funcionam como um 'seguro' para as distribuidoras, podendo ser vistas também como mais um fator de incremento à confiabilidade.

O consumidor paga ao agente distribuidor pela garantia da disponibilidade da demanda contratada, estando este sujeito a penalidades severas, caso não disponibilize este nível de potência quando solicitado. Como por sua vez, o agente distribuidor adquire dos agentes geradores a energia que distribui, transfere a estes a responsabilidade de tornar disponível o montante de demanda de seus contratos com consumidores cativos. Por este 'serviço', os agentes geradores são remunerados através das cláusulas de potência dos contratos iniciais e bilaterais, estando também sujeitos a penalidades, se não honrarem tal responsabilidade. No caso dos consumidores livres, estes em seus contratos, também exigem dos agentes geradores a garantia de um certo nível de demanda, ou seja, de uma certa capacidade de geração, pagando-os por isto.

Pode-se dizer então que existe uma 'transferência remunerada' de responsabilidades, no sentido consumidor-gerador, passando pelos agentes distribuidores. Este mecanismo representaria então uma outra maneira de proporcionar os sinais econômicos necessários à expansão da capacidade de geração do sistema, uma vez que os agentes geradores não se sujeitarão a serem pesadamente penalizados, ao não terem como proporcionar a capacidade de geração exigida pelos distribuidores e consumidores livres

contratados. Haverá novos investimentos em capacidade a partir do momento em que o risco de não se conseguir atender à demanda requerida estiver acima dos padrões considerados ideais.

2.8 Conclusões

2.8.1 Aplicação dos Encargos de Capacidade

Apesar de não ser objetivo deste trabalho, é importante que se façam algumas considerações sobre a necessidade e propriedade de se aplicarem os encargos de capacidade, principalmente no sistema elétrico brasileiro, segundo o modelo proposto em [MAE, 2000a, 2000b; CRMSE, 2002a].

Em qualquer sistema em que as tarifas de energia elétrica possuam componente de demanda, deve-se ter um cuidado extra, caso se decida também pela utilização dos encargos de capacidade, na tentativa de se diminuir a volatilidade do preço do mercado a vista e incentivar a expansão da capacidade instalada, principalmente em relação à capacidade despachada. O grande problema que poderia ser criado seria a duplicidade da remuneração dos agentes geradores por capacidade, pois a mesma já estaria contemplada nas tarifas, pelo menos em parte. No caso brasileiro, apenas os consumidores residenciais e pequenos consumidores comerciais e industriais não pagam por capacidade, por não existir ainda um sistema de medição em tempo real que permita o acesso à demanda individual de toda a carga.

As metodologias apresentadas no Apêndice C propõem que os encargos de capacidade, referentes à potência despachada, sejam adicionados ao preço do MAE. Entretanto, apenas uma parte de toda a energia comercializada, a energia não contratada, tanto através dos contratos iniciais quanto dos contratos bilaterais, seria negociada no mercado atacadista. Haveria, portanto, um confinamento de tais encargos, pois apenas os distribuidores, comercializadores e consumidores livres que recorressem ao MAE, arcariam com este custo. Mecanismos deveriam ser previstos, para evitar que ele fosse repassado à carga pelos distribuidores, o que representaria um ônus desnecessário aos consumidores, uma vez que estes já os teriam pago nas tarifas, principalmente aqueles

que possuem maior custo de interrupção, tais como consumidores industriais e comerciais de maior porte.

Quanto à capacidade não despachada, conforme relatado na Seção 2.7.2, ela também já estaria sendo remunerada, uma vez que itens de demanda dos contratos entre consumidores e distribuidores e entre estes e os geradores, definem valores de demanda que, mesmo se não utilizados serão pagos. Os distribuidores, apesar de atenderem cargas sem estes contratos, garantem junto aos geradores a disponibilidade para este atendimento.

Sob este aspecto, a diferença básica entre as duas metodologias é que a primeira considera a influência dos contratos iniciais, aceitando que eles já teriam contemplado 80% dos encargos por capacidade, o que muito provavelmente se manteria, com a gradual extinção dos contratos iniciais e a sua natural substituição pelos contratos bilaterais. Apesar de não haver uma certeza a respeito deste percentual, a metodologia cujo resumo é apresentado na Seção C.1 [MAE, 2000a; MAE, 2000b] pode ser considerada mais coerente com a proposta de se garantir uma remuneração regular para as usinas de ponta e uma diminuição dos preços do MAE, já que elas geralmente representam apenas uma parcela da disponibilidade não despachada.

Na verdade, no Brasil, o aumento da presença destas usinas, geralmente térmicas, raramente despachadas, ou despachadas apenas em situações emergenciais, o que ajudaria a justificar o uso dos encargos de capacidade, não se confirmou, continuando insignificante a sua quantidade. Devido à sua característica particular, no sistema brasileiro as usinas térmicas são quase sempre usinas de base e a caracterização das usinas de ponta se dá por questões hidrológicas, diferentemente dos sistemas com predominância de geração térmica.

Observa-se que a metodologia da Seção C.1 apresenta bastante afinidade com a proposta de remuneração em separado das reservas operativas e das reservas de planejamento representada na Figura 2.2, pelo fato de que, em ambos os casos, os encargos atingiriam apenas parte da capacidade não despachada. Assim ela poderia ser utilizada para se definir a remuneração das reservas de planejamento, incluindo as reservas operativas de substituição R4 (ver Seção A.2), que lhes são correlatas, uma vez estabelecidos os seus percentuais, em relação à capacidade disponível total.

Esta remuneração já seria suficiente para cobrir os custos fixos das usinas de ponta não despachadas, sendo ela repassada a toda a carga. Como isto representa uma aquisição antecipada de energia por toda a carga, seria necessária a criação de um procedimento anual específico de contabilização, que determinasse quais as usinas, caracterizadas realmente como 'de ponta', tivessem sido novamente remuneradas, por terem sido despachadas, e em qual montante. A remuneração em excesso seria ressarcida à carga através de créditos futuros de consumo.

Apesar de muito estudado e, por duas vezes cogitada a sua implementação, até o momento não houve uma ação efetiva para que os encargos por capacidade fossem adotados no sistema elétrico brasileiro. Na nova proposta de modelo para o setor elétrico brasileiro, a ser operacionalizado a partir de 2004, ainda não é possível observar-se qualquer referência a uma possível aplicação destes encargos. De toda forma, uma decisão a seu respeito deveria ser pautada na sua real necessidade, nas possíveis interferências com outras modalidades de encargos já estabelecidos e na experiência adquirida, que vem mostrando que eles não alcançaram os objetivos a que se propuseram, além de permitirem fraudes e manipulações de preços.

2.8.2 Utilização dos Encargos de Capacidade da Disponibilidade não Despachada como Remuneração das Reservas de Potência Operativa.

Na Seção 2.6.2 procurou-se justificar, sob o aspecto técnico, o fato de que as reservas de planejamento e as reservas de potência operativa devem ser consideradas e tratadas como dois serviços com características próprias, que os diferenciam grandemente.

Sob o ponto de vista econômico, é necessário ficar claro que as suas remunerações devem ser também diferenciadas. Conforme se observa em diversos sistemas elétricos, com o objetivo de garantir a confiabilidade, geralmente adota-se uma política relacionada às reservas operativas ou uma política relacionada aos pagamentos por capacidade. Eventualmente as duas são adotadas, sendo neste caso, remuneradas separadamente, distinguindo-se claramente cada parcela de reserva.

Não é aceitável considerar-se que os encargos de capacidade já proporcionem remuneração também para as reservas operativas para confiabilidade e regulação, o que se justifica por todas as considerações já feitas no Capítulo 2, resumidas a seguir:

- os encargos de capacidade referentes à disponibilidade não despachada, que na verdade representam pagamentos por capacidade, possuem um perfil de recuperação de custos fixos de investimentos em geração na comercialização da energia, ao passo que as reservas operativas possuem um perfil de serviço prestado, envolvendo custos completamente distintos daqueles;
- o custo marginal relacionado às reservas operativas é o de curto prazo e às reservas de planejamento é o de longo prazo, que envolve o custo de expansão do sistema;
- os serviços de reservas operativas e regulação de frequência possuem um valor econômico, pelo benefício que trazem ao sistema, que precisa ser remunerado, como qualquer outro serviço, sendo o benefício e o valor econômico das reservas de planejamento muito diferentes;
- muitas vezes os serviços de reservas operativas e regulação de frequência exigem instalações e equipamentos específicos, gerando custos de investimentos extras, independentes daqueles envolvidos na expansão da geração, além de eventual acréscimo de outros custos fixos e variáveis;
- os pagamentos por capacidade podem ser interpretados como um contrato de venda de energia, no qual a definição do preço e o pagamento da energia é feito de forma antecipada, sendo a energia entregue no momento em que vier a ocorrer um corte de carga no sistema, de forma absolutamente diversa do que se propõe para os serviços de reservas operativas e regulação de frequência;
- considerando-se apenas o serviço de regulação de frequência, uma vez que o seu objetivo não se relaciona diretamente com os requisitos de confiabilidade, mas principalmente com a qualidade da energia produzida, ele de forma alguma, deve ter a sua remuneração relacionada aos pagamentos por capacidade.

Na Figura 2.2 apresentou-se uma proposta de remuneração em separado das reservas operativas e das reservas de planejamento. Estas seriam remuneradas, por exemplo, por meio dos encargos de capacidade relativos à disponibilidade não despachada, e as reservas operativas por meio de uma remuneração específica. Como se observa, apenas parte das reservas de planejamento seria alvo de remuneração própria, exatamente aquela disponível, referente às usinas de ponta do sistema, que se confunde com a parcela das reservas operativas conhecidas como reservas de substituição. Desta forma, soluciona-se a questão do pagamento das reservas de substituição, além da questão específica das usinas de ponta, quase sempre não despachadas, que teriam a cobertura de seus custos fixos feita de forma mais regular. As demais usinas, que poderiam ou não encontrarem-se despachadas, poderiam ter a contribuição da sua potência não despachada à confiabilidade do sistema, remunerada na forma de reserva operativa.