EVO

EFFICIENCY VALUATION ORGANIZATION

Protocolo Internacional de Medição e Verificação de *Performance*

Estatística e Incerteza para o PIMVP

Preparado pela Efficiency Valuation Organization
www.evo-world.org

Junho 2014

EVO 10100 – 1:2014 (Br)

Visão da EVO

Um mercado global que valorize corretamente o uso eficiente dos recursos naturais e utilize as opções de uso final eficiente como uma alternativa viável às opções de suprimento.

Missão da EVO

Desenvolver e promover o uso de protocolos normatizados, métodos e ferramentas para quantificar e gerenciar os riscos de desempenho e benefícios associados com transações comerciais com eficiência energética nos usos finais, energias renováveis e eficiência no uso da água.

Sumário Página

1 Introdução

O objetivo da M&V é determinar a economia de energia de forma confiável. Para que a determinação da energia seja confiável, é preciso que haja um nível de incerteza razoável. A incerteza da economia pode ser gerenciada controlando-se os erros aleatórios e sistemáticos. Os erros aleatórios são afetados pela qualidade dos equipamentos de medição, pelas técnicas de medição e pela concepção do procedimento de medição. Os erros sistemáticos são afetados pela qualidade dos dados de medição, suposições e análises. Reduzir erros significa, em geral, aumentar custos – destarte, a necessidade de uma melhor incerteza deve ser justificada pelo valor da melhoria da informação.

O cálculo da economia de energia envolve a comparação de dados medidos de energia e um cálculo de "ajustes" para converter ambas as medições para o mesmo conjunto de condições de operação. Tanto as medições quanto os ajustes introduzem erros. Erros podem aparecer devido à imprecisão do medidor, procedimentos de amostragem ou de ajuste. Estes processos produzem "estimativas" estatísticas com valores determinados ou esperados e algum nível de variação. Em outras palavras, os valores verdadeiros não são conhecidos, apenas estimativas com algum nível de incerteza. Todas as medições físicas e análises estatísticas são baseadas em estimativas de tendências centrais, como valores médios, e quantificação da variação como faixa de variação, desvio padrão, erro padrão e variância.

Estatística é o corpo de métodos matemáticos que podem ser aplicados a dados para ajudar a tomar decisões em face da incerteza. A estatística fornece os meios para verificar se os resultados determinados são "significativos", ou seja, que seja mais provável que sejam um efeito real da AEE (ação de eficiência energética) do que um comportamento aleatório.

Erros podem ocorrer de três maneiras: modelagem, amostragem e medição.

- **Modelagem**. Erros na modelagem matemática devido a forma funcional não apropriada, inclusão de variável irrelevante, exclusão de variável relevante, etc.
- **Amostragem**. Erros de amostragem ocorrem quando somente uma porção da população dos valores reais é medida, ou uma amostra enviesada é usada. A representação de somente uma parte da população pode ocorrer em um sentido físico ou de tempo.
- **Medição**. Erros de medição surgem da imprecisão de sensores, rastreamento do medidor, perda de precisão desde a calibração, medições imprecisas, etc. A magnitude de tais erros é principalmente dada pela especificação do fabricante e gerenciada por recalibração periódica.

Este documento orienta quanto à quantificação das incertezas criadas por estas três formas de erros. Algumas fontes de erros são desconhecidas e não quantificáveis:

- ➢ Seleção ou colocação inadequada de medidores
- ➢ Estimativas imprecisas na Opção A, ou
- ➢ Estimativas inadequadas de efeitos interativos nas Opções A e B.

Incertezas não conhecidas ou não quantificáveis podem somente ser gerenciadas seguindo-se as melhores práticas da indústria.

1.1 Expressão da incerteza

Para se expressar a economia de uma forma estatisticamente válida, a economia necessita ser expressa em conjunto com seus níveis de precisão e confiança. Confiança refere-se à probabilidade de que a economia estimada fique dentro da faixa de precisão[1].

EXEMPLO O processo de estimar a economia pode levar a uma afirmação tal como: "a melhor estimativa da economia é 1.000 kWh anuais (estimativa pontual) com 90% de probabilidade (confiança) de que o valor da economia real média fique dentro de ±20% de 1.000". Uma representação gráfica desta relação é mostrada na Figura 1.

[1] Termos estatísticos selecionados estão definidos na Seção 1.3.

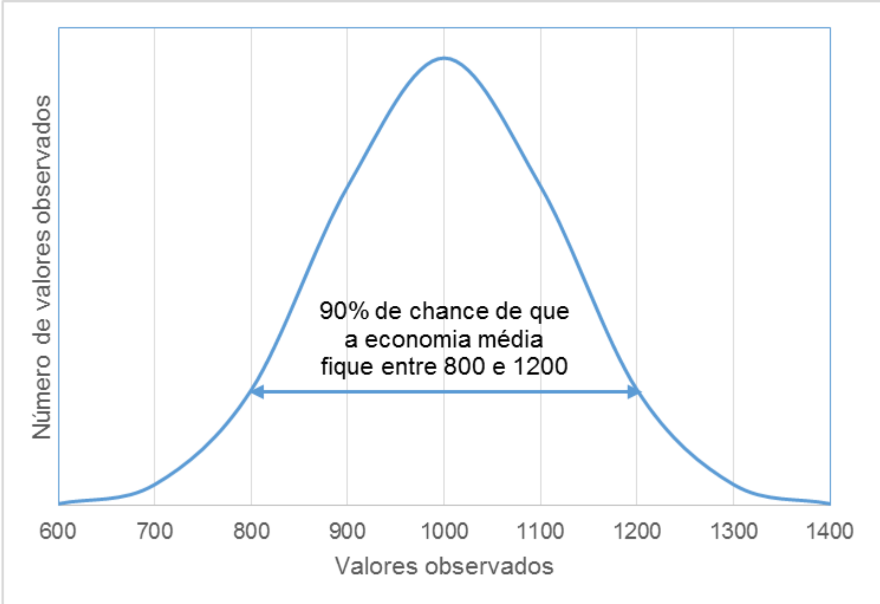

Figura 1 Distribuição normal da população

Uma declaração da precisão estatística (os ±20%) sem um intervalo de confiança (os 90%) é imprecisa. O processo de M&V pode produzir uma precisão extremamente alta, porém com baixo nível de confiança (o que não aumenta a confiabilidade do processo).

Exemplo: A economia pode ser estabelecida com uma precisão de ±1%, porém o nível de confiança pode cair de 95% para 35%.

1.2 Incerteza aceitável

A economia é considerada estatisticamente válida se for relativamente grande em relação às variações estatísticas. Especificamente, a economia deve ser maior que duas vezes o erro padrão do valor da linha de base. Se a variância da linha de base for muito grande, o comportamento aleatório não explicado no uso da energia da instalação ou sistema é grande e uma determinação simples da economia não é confiável.

Quando não se conseguir atender a este critério (economia > 2 x erro padrão), considerar usar:

- Instrumentos de medição mais precisos,
- Mais variáveis independentes no modelo matemático,
- Maiores tamanhos de amostras, ou
- Uma Opção do PIMVP que seja menos afetada pelas variáveis não conhecidas.

1.3 Definições de termos estatísticos

Média amostral ($\bar{Y}$): determinada pela soma dos pontos individuais (Y_i) dividida pelo número total de pontos (n), como segue:

$$\bar{Y} = \frac{\sum Y_i}{n} \qquad (1)$$

Variância amostral (S^2): mede a extensão que os valores observados diferem uns dos outros, ou seja, sua variabilidade ou dispersão. Quanto maior a variabilidade, maior a incerteza da média. É calculada pela média dos quadrados dos desvios individuais em relação à média. A razão de usar os quadrados dos desvios é simplesmente para evitar valores negativos (quando um valor está abaixo da média), que cancelariam os valores positivos (acima da média). É calculada como segue:

$$S^2 = \frac{\sum (Y_i - \bar{Y})^2}{n-1} \qquad (2)$$

Desvio padrão amostral (s): é simplesmente a raiz quadrada da variância amostral. Isto traz a medida da variabilidade para a mesma dimensão dos dados (ou seja, se a unidade de variância é kWh^2, a unidade do desvio padrão é kWh, a mesma dos dados originais).

$$s = \sqrt{S^2} \qquad (3)$$

Erro padrão da amostra (EP): é o desvio padrão amostral dividido por $\sqrt{n}$. Esta medida é usada para estimar a precisão da média amostral. Também é simbolizado por $\bar{s}$, ou o "desvio padrão amostral da média" em muitos livros de estatística.

$$EP = \frac{s}{\sqrt{n}} \qquad (4)$$

Desvio padrão amostral do total (s_{tot}): muitas vezes o interesse está nas propriedades estatísticas do total em vez da média. O desvio padrão amostral do total é usado para definir a precisão do total da amostra. É definido como a raiz quadrada do tamanho da amostra ($\sqrt{n}$) vezes o desvio padrão amostral.

$$s_{tot} = \sqrt{n} \bullet s \qquad (5)$$

Coeficiente de variação (cv): é simplesmente o desvio padrão de uma distribuição, expresso como uma porcentagem da média. Por exemplo, o *cv* de uma amostra total é o s_{tot} dividido pela soma total dos valores da amostra; o *cv* de uma amostras de médias seria o erro padrão dividido pela média da amostra. A fórmula geral é:

$$cv = \frac{s}{\bar{Y}} \qquad (6)$$

Precisão: é a medida absoluta ou relativa da faixa dentro da qual o valor verdadeiro é esperado ocorrer com algum nível de confiança especificado. O nível de confiança refere-se à probabilidade que a citada faixa contenha o parâmetro estimado.

Precisão absoluta: é calculada a partir do erro padrão da amostra usando-se a estatística *t* de Student (ver a Tabela 1). Além da Tabela 1, a distribuição *t* pode ser achada em tabelas estatísticas, livros, internet ou *softwares*.

$$I_a = t \times EP_{\bar{Y}} \text{ ....} \qquad (7)$$

Tabela 1 – Distribuição t de Student

Graus de liberdade	**Nível de confiança**				**Graus de liberdade**	**Nível de confiança**			
GL	**95%**	**90%**	**80%**	**50%**	**GL**	**95%**	**90%**	**80%**	**50%**
1	12,71	6,31	3,08	1,00	16	2,12	1,75	1,34	0,69
2	4,30	2,92	1,89	0,82	17	2,11	1,74	1,33	0,69
3	3,18	2,35	1,64	0,76	18	2,10	1,73	1,33	0,69
4	2,78	2,13	1,53	0,74	19	2,09	1,73	1,33	0,69
5	2,57	2,02	1,48	0,73	21	2,08	1,72	1,32	0,69
6	2,45	1,94	1,44	0,72	23	2,07	1,71	1,32	0,69
7	2,36	1,89	1,41	0,71	25	2,06	1,71	1,32	0,68
8	2,31	1,86	1,40	0,71	27	2,05	1,70	1,31	0,68

Graus de liberdade GL	Nível de confiança 95%	90%	80%	50%	Graus de liberdade GL	Nível de confiança 95%	90%	80%	50%
9	2,26	1,83	1,38	0,70	31	2,04	1,70	1,31	0,68
10	2,23	1,81	1,37	0,70	35	2,03	1,69	1,31	0,68
11	2,20	1,80	1,36	0,70	41	2,02	1,68	1,30	0,68
12	2,18	1,78	1,36	0,70	49	2,01	1,68	1,30	0,68
13	2,16	1,77	1,35	0,69	60	2,00	1,67	1,30	0,68
14	2,14	1,76	1,35	0,69	120	1,98	1,66	1,29	0,68
15	2,13	1,75	1,34	0,69	☐	1,96	1,64	1,28	0,67

Nota: Calcular os GL da forma abaixo:

- GL = n – 1 (para uma distribuição amostral)
- GL = n – p – 1 (para um modelo de regressão)

Onde:
n = tamanho da amostra
p = número de variáveis independentes no modelo de regressão

Em geral, o valor verdadeiro de qualquer estimativa estatística é esperado ficar, dentro de um dado nível de confiança, na faixa definida por

Faixa = Estimativa ± precisão absoluta ....(8)

Onde "estimativa" é qualquer valor empiricamente derivado de um parâmetro de interesse (por exemplo, consumo total, número médio de unidades produzidas, etc.).

Precisão relativa é a precisão absoluta dividida pela estimativa:

$$I_{rel} = \frac{t \bullet EP}{Estimativa} \qquad (9)$$

Exemplo: Considere os dados da Tabela 2 de 12 leituras mensais de um medidor, e a respectiva análise da diferença entre as leituras.

Tabela 2 – Exemplo de análise de dados

Mês	Leitura real	Diferenças da média Direta	Diferenças da média Quadrado
1	950	-50	2.500
2	1.090	90	8.100
3	850	-150	22.500
4	920	-80	6.400
5	1.120	120	14.400
6	820	-180	32.400
7	760	-240	57.600
8	1.210	210	44.100
9	1.040	40	1.600
10	930	-70	4.900
11	1.110	110	12.100
12	1.200	200	40.000
Total	12.000	0	246.600

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{12.000}{12} = 1.000$$

O valor médio é:

A variância é: $S^2 = \frac{\sum (Y_i - \bar{Y})^2}{n-1} = \frac{246.600}{12-1} = 22.418$

O desvio padrão é: $s = \sqrt{S^2} = \sqrt{22.418} = 150$

O erro padrão é: $EP = \frac{s}{\sqrt{n}} = \frac{150}{\sqrt{12}} = 43$

Na Tabela 2 há 12 pontos. Significa que GL = 12 – 1 = 11. Usando-se a Tabela 1, para um nível de confiança de 90%, o valor de *"t"* é 1,80. Desta forma

A precisão absoluta é: $t \bullet SE = 1{,}80 \times 43 = 77$

E a precisão relativa é: $\frac{t \bullet EP}{estimativa} = \frac{77}{1.000} = 7{,}7\%$

Então, há 90% de confiança que o valor mensal verdadeiro do consumo fique na faixa entre 923 e 1.077 kWh. Pode-se dizer com 90% de confiança que o valor médio de 12 observações é 1.000 ± 7,7%. Da mesma forma, pode-se dizer:

- Com 95% de confiança que o valor médio das 12 observações é 1.000 ± 9,5% ou
- Com 80% de confiança que o valor médio das 12 observações é 1.000 ± 5,8% ou
- Com 50% de confiança que o valor médio das 12 observações é 1.000 ± 3,0%.

2 Modelagem

Modelos matemáricos são usados em M&V para preparar o termo referente aos ajustes de rotina nas várias versões da equação de cálculo da economia de energia discutidas no "Conceitos Básicos do PIMVP", abaixo transcritas:

1. Economia = (Consumo ou demanda da linha de base – Consumo ou demanda do período de determinação da economia) ± Ajustes
2. Economia = (Energia da linha de base – Energia do período de determinação) ± Ajustes de rotina ± Ajustes não de rotina
3. Energia evitada (ou Economia) = (Energia da linha de base ± Ajustes de rotina às condições do período de determinação ± Ajustes não de rotina às condições do período de terminação) – Energia do período de determinação
4. Energia evitada (ou Economia) = Energia da linha de base ajustada – Energia do período de determinação da economia ± Ajustes não de rotina da energia da linha de base às condições do período de determinação da economia
5. Economia normalizada = (Energia da linha de base ± Ajustes de rotina às condições fixas ± Ajustes não de rotina às condições fixas) – (Energia do período de determinação ± Ajustes de rotina às condições fixas ± Ajustes não de rotina às condições fixas)
6. Economia na Opção A = Valor estimado x (Parâmetro medido na linha de base – Parâmetro medido no período de determinação da economia)
7. Economia na Opção B = Energia da linha de base – Energia do período de determinação da economia
8. Economia = Energia da linha de base do modelo calibrado [hipotético ou sem AEEs] – Energia do período de determinação da economia do modelo calibrado [com AEEs]
9. Economia = Energia da linha de base do modelo calibrado [hipotético ou sem AEEs] – Energia medida no período de calibração ± Erro de calibração na respectiva leitura

A ação de modelagem envolve obter uma relação matemática entre as variáveis dependentes e indenpendentes. A variável dependente, usualmente a energia, é modelada como sendo regida por uma ou mais variáveis independentes X_i (também conhecidas como "variáveis explicativas"). Este tipo de modelagem é chamado de análise de regressão.

Na análise de regressão, o modelo tenta "explicar" a variação na energia resultante das variações nas variáveis independentes individuais (X_i). Por exemplo, se uma das *Xs* é a taxa de produção, o modelo analisará se a variação da energia em torno da sua média é causada por mudanças na taxa de produção. O modelo quantifica o quanto. Por exemplo, quando a produção cresce em uma unidade, o consumo de energia cresce "*b*" unidades, onde "*b*" é chamado de coeficiente de regressão.

O modelo mais comum é o de regressão linear na forma:

$$Y = b_o + b_1X_1 + b_2X_2 + \ldots.. + b_pX_p + e$$

Onde:

- Y é a variável dependente, usualmente na forma de consumo de energia em períodos específicos (por exemplo, 30 dias, 1 semana, 1 dia, 1 hora, etc.)
- X_i ($i = 1, 2, 3, \ldots p$) representam as 'p' variáveis independentes como clima, produção, ocupação, período de medição, etc.
- b_i ($i = 0, 1, 2, \ldots p$) representam os coeficientes derivados para cada variável independente, e um coeficiente fixo (b_0) não relacionado às variáveis independentes
- e representa o erro residual que permanece inexplicado depois de se levar em conta o impacto das várias variáveis independentes. A análise de regressão encontra o conjunto de valores b_i que minimizam a soma dos quadrados dos erros residuais (assim, o modelo de regressão também é chamado de modelo dos mínimos quadrados).

Um exemplo do modelo acima para o consumo de energia de um prédio é:

Consumo mensal de energia = 342.000 + (63 x *GDA*) + (103 x *GDR*) + (222 x *Ocupação*)

GDA e GDR são os graus dia de aquecimento (GDA) e refrigeração (GDR). A ocupação é a medida em percentual de ocupação do prédio. Neste modelo, 342.000 é uma estimativa da carga base em kWh, 63 mede a mudança em consumo para cada GDA adicional, 103 mede a mudança no consumo para cada GDR adicional e 222 mede a mudança no consumo para cada 1% de alteração na ocupação.

2.1 Erros de modelagem

Ao usar modelos de regressão, como descrito acima, muitos tipos de erros podem ser introduzidos, como descrito abaixo.

1. O modelo é construído com valores fora da faixa provável da variável. O modelo matemático deve ser construído somente usando-se valores razoáveis das variáveis dependentes e indepedentes.
2. O modelo matemático pode não incluir variáveis independentes relevantes, introduzindo a possibilidade de relações tendenciosas (chamada de viés de variável omitida).
3. O modelo pode incluir algumas variáveis irrelevantes.
4. O modelo pode usar uma forma funcional não apropriada.
5. O modelo pode ser baseado em dados não suficientes ou não representativos.

Esses erros são discutidos em maiores detalhes abaixo.

2.1.1 Uso de dados fora do intervalo

Se o modelo for concebido com dados que não são representativos do comportamento energético normal da instalação, as predições podem não ser confiáveis. Isto pode incluir a inclusão de pontos fora da curva (*outliers*) ou valores que estejam fora da faixa de razoabilidade. Os dados devem ser verificados quanto a representação do comportamento energético real da instalação antes de conceber o modelo.

2.1.2 Omissão de variáveis relevantes

No processo da M&V, a análise de regressão é usada para dar conta das variações do consumo de energia. Muitos sistemas energéticos são afetados por inúmeras variáveis independentes. Não se pode esperar que os modelos de regressão incluam todas as variáveis independentes. Mesmo se isso fosse possível, o modelo seria demasiadamente complicado para ser útil e requeriria um esforço excessivo para coletar todos os dados necessários. A forma prática de abordar este problema é incluir apenas as variáveis que tenham um impacto significativo na variação da energia.

No entanto, a omissão de uma variável independente relevante pode ser um erro crucial. Se uma variável independente não é considerada (por exemplo, GDA, produção ou ocupação), o modelo não conseguirá explicar uma parte importante da variação da energia. Este modelo deficiente irá também atribuir alguma variação, que na verdade é devida à variável omitida, às outras variáveis incluídas no modelo. O efeito geral será um modelo menos preciso.

Não há indicações óbvias deste problema nos testes estatísticos padrão (exceto talvez um baixo R^2). A experiência e conhecimento da engenharia energética do sistema cujo desempenho está sendo medido são muito importantes para tratar esta questão.

Pode haver casos onde se saiba que existe uma relação com uma variável registrada durante o período da linha de base mas que, devido à falta de orçamento para incluir a sua coleta de dados durante o período de determinação da economia, não foi incluída no modelo da linha de base. Tal omissão de variável relevante deve ser registrada e justificada no Plano de M&V.

2.1.3 Inclusão de variáveis irrelevantes

Os modelos podem eventualmente incluir variáveis independentes irrelevantes. Se esta variável irrelevante não tiver relação (correlação) com as variáveis relevantes incluídas no modelo, ela terá um impacto mínimo. No entanto, se a variável irrelevante for correlacionada com outras variáveis relevantes no modelo, poderá enviesar os coeficientes das variáveis relevantes.

É preciso cuidado ao adicionar mais variáveis independentes em uma análise de regressão somente porque estão disponíveis. Avaliar a relevância de variáveis independentes requer tanto experiência quanto intuição. A estatística *t* é uma forma de confirmar a relevância de uma variável independente. A experiência na análise do comportamento da energia no tipo de instalação em questão em qualquer programa de M&V é necessária para determinar a relevância de variáveis independentes.

2.1.4 Forma funcional

Muitas vezes é possível modelar a relação entre as variáveis dependentes e independentes com uma forma funcional incorreta. Por exemplo, uma relação linear pode ser incorretamente usada para modelar uma relação física subjacente que não seja linear. Por exemplo, o consumo de energia elétrica e a temperatura ambiente tendem a possuir uma relação não linear (frequentemente em forma de ‘U’) em edifícios com refrigeração e aquecimento ambientais a eletricidade, quando se considera um ano de funcionamento (o consumo de energia elétrica é alto no verão e inverno, e relativamente baixo nas meias-estações). Fazer um modelo linear desta relação introduz erros desnecessários. Ao contrário, modelos separados lineares para cada estação devem ser usados.

Também pode ser apropriado tentar relações de ordem mais elevada, por exemplo, $Y = f(X, X^2, X^3)$.

O modelador precisa avaliar diferentes formas funcionais e selecionar a mais apropriada usando parâmetros de avaliação.

2.1.5 Falta de dados

Podem também ocorrer erros provenientes de insuficiência de dados, tanto em termos de quantidade (ou seja, poucos pontos de medição) quanto de período de medição (por exemplo usar somente medições no verão e tentar extrapolar para o inverno). Os dados utilizados para modelar a energia devem ser representativos da faixa de operação da instalação ou sistema. O período de medição para elaborar o modelo deve incluir várias possíveis estações, formas de uso da energia, etc. Isto pode incluir uma extensão do período de medição ou dos tamanhos das amostras.

2.2 Avaliação dos modelos de regressão

Para avaliar quão bem um determinado modelo de regressão explica a relação entre o uso da energia e as variáveis independentes, três testes são mais usados, como descrito abaixo.

2.2.1 Coeficiente de determinação (R^2)

O primeiro passo que deve ser dado para avaliar a precisão de um modelo é examinar o coeficiente de determinação R^2, que é uma medida do quanto da variação da energia é explicado pelo modelo (ou seja, quanto da variação da variável dependente Y em relação ao seu valor médio é explicado pelo modelo de regressão). Matematicamente, o R^2 é igual a:

$$R^2 = \frac{Variação\ explicada\ em\ Y}{Variação\ total\ em\ Y}$$

Ou mais explicitamente:

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \qquad (10)$$

Onde:

- $\hat{Y}_i$: valor da energia previsto pelo modelo em um ponto dado pela medição do valor da variável independente (i.e., obtido pela inserção dos valores de *X* no modelo de regressão)
- $\bar{Y}$: média nos *n* valores medidos da energia, calculados pela equação (1).
- Y_i : valor real medido da energia

Todos os pacotes estatísticos e ferramentas de análise de regressão com planilhas calculam o valor de R^2.

A faixa de variação possível para o R^2 é de zero a um. Um R^2 igual a zero significa que nenhuma variação da energia é explicada pelo modelo, desta forma o modelo não ajuda a entender as variações em *Y* (isto é, as variáveis independentes selecionadas não fornecem qualquer explicação das causas das variações observadas em *Y*). Por outro lado, um R^2 igual a um significa que o modelo explica 100% das variações em *Y* (isto é, o modelo prediz *Y* com total certeza, para cada conjunto de valores das variáveis independentes). Nenhum destes dois valores limites (0 e 1) ocorre na prática.

De forma geral, quanto maior for o coeficiente de determinação melhor a descrição do modelo da relação entre as variáveis dependentes e independentes. Embora não haja um padrão universal para um valor mínimo de aceitação de R^2, 0,75 é usualmente considerado um indicador mínimo razoável de uma boa relação causal entre a energia e as variáveis independentes.

O teste do R^2 deve somente ser usado como uma verificação inicial. Os modelos não devem ser aceitos ou rejeitados somente com base no R^2. Finalmente, um R^2 baixo é uma indicação que alguma variável ou variáveis não estão incluídas no modelo, ou que a forma funcional do modelo (por exemplo, linear) não é apropriada. Nesta situação, deve-se considerar variáveis independentes adicionais ou tentar diferentes formas funcionais.

2.2.2 Erro padrão da estimativa

Quando um modelo é usado para predizer um valor de energia (*Y*) para um conjunto de variáveis independentes, a precisão da predição é medida pelo erro padrão da estimativa ($EP_{\hat{Y}}$). Esta medida da precisão é fornecida por todos os pacotes padrão e planilhas de regressão.

Uma vez que se insere um conjunto de valores das variáveis independentes no modelo de regressão para estimar um valor de energia ($\hat{Y}$), uma aproximação da faixa de possíveis valores de $\hat{Y}$ pode ser calculada usando-se a equação 8 como:

$$\hat{Y} \pm t \times EP_{\hat{Y}}$$

onde:

- $\hat{Y}$ é o valor de energia (*Y*) predito pelo modelo de regressão
- *t* é o valor obtido da estatística *t* de Student (ver a Tabela 1)
- $EP_{\hat{Y}}$ é o erro padrão da estimativa (predição). É calculado como:

$$EP_{\hat{Y}} = \sqrt{\frac{\sum (\hat{Y}_i - Y_i)^2}{n - p - 1}} \qquad (11)$$

Onde *p* é o número de variáveis independentes na equação de regressão.

Esta estatística é usualmente referida como erro médio quadrático (*EMQ*). Dividindo-se o *EMQ* pela média do consumo de energia tem-se o coeficiente de variação do *EMQ*, ou *CV(EMQ).*

$$CV(EMQ) = \frac{EP_{\hat{Y}}}{\overline{Y}} \qquad (12)$$

Uma medida similar é o erro médio sistemático (*EMS*), definido como:

$$EMS = \frac{\sum(\hat{Y}_i - Y_i)}{n} \qquad (13)$$

O *EMS* é um bom indicador do viés geral na estimativa de regressão. Um *EMS* positivo indica que as estimativas da regressão tendem a superestimar os valores reais. O total de vieses positivos tendem a serem cancelados pelos negativos. O *EMQ* não tem este problema de cancelamento.

Todas as três medidas podem ser usadas para avaliar a calibração de modelos de simulação da Opção D.

2.2.3 Estatística t

Já que os coeficientes do modelo de regressão (b_k) são estimativas estatísticas da verdadeira relação entre cada variável *X* e *Y*, são sujeitos a variação. A precisão da estimativa é medida pelo erro padrão do coeficiente e do valor respectivo da estatística *t*. A estatística *t* é um teste estatístico para determinar se uma estimativa tem significado estatístico. Uma vez que o valor é estimado usando-se o teste, pode ser comparado com os valores críticos *t* da Tabela 1.

O erro padrão de cada coeficiente é calculado por qualquer *software* de regressão. A equação abaixo se aplica para o caso de uma variável independente.

$$EP_b = \sqrt{\frac{\sum(Y_i - \hat{Y})^2 / (n-2)}{\sum(X_i - \bar{X})^2}} \qquad (14)$$

Para os casos com mais de uma variável independente, a equação fornece uma aproximação razoável quando as variáveis independentes o são realmente (ou seja, não são correlacionadas). Se não, a equação fica muito complexa e o analista de M&V deve utilizar um pacote de *software* de regressão para calcular os erros padrão dos coeficientes. A faixa dentro da qual o valor verdadeiro do coeficiente deve estar é encontrada pela equação (8):

$$b \pm t \times EP_b$$

O erro padrão do coeficiente *b* leva também ao cálculo da estatística *t*. Este teste, em última instância, determina se o coeficiente calculado tem significado estatístico. A estatística *t* é calculada por todos os *softwares* estatísticos usando a seguinte equação:

$$\text{Estatística } t = \frac{b}{SE_b} \qquad (15)$$

Após a estimativa da estatística *t* pela equação acima, o valor pode ser comparado aos valores críticos da Tabela 1. Se o valor absoluto da estatística *t* calculada exceder o respectivo valor da Tabela 1, pode-se concluir que a estimativa do coeficiente é estatísticamente válida.

Uma regra empírica estabelece que se o valor absoluto da estatística *t* for maior que 2 significa que o coeficiente estimado é significativo em relação ao erro padrão, ou seja, há um relação válida entre *Y* e o *X* do respectivo coeficiente; ou ainda, pode-se concluir que o *b* estimado não é zero. Entretanto, uma estatística *t* perto de 2 significa que a precisão do valor do coeficiente está em torno de ±100%, o que não é muito para se confiar no valor de *b* calculado. Para se obter uma melhor precisão, digamos ±10%, os valores da estatística *t* devem ficar em torno de 20, ou o erro padrão de *b* não ser maior que 10% do seu valor.

Para melhorar o resultado da estatística *t*, considerar as seguintes ações:

- Selecionar as variáveis independentes com a relação mais forte com a energia;
- Selecionar as variáveis independentes cuja faixa de variação seja a maior possível (se *X* não varia muito no modelo de regressão, não se pode estimar bem o valor de *b* e a estatística *t* será pobre);
- Coletar e usar mais pontos de dados ao desenvolver o modelo; ou
- Selecionar outra forma funcional para o modelo; por exemplo, uma que separe os coeficientes para cada estação em um prédio cuja energia esteja bastante afetada pelas mudanças sasonais.

3 Amostragem

A amostragem cria erros porque nem todas as unidades sob estudo são medidas. A situação mais simples de amostragem é selecionar aleatoriamente *n* unidades em uma população total de *N* unidades. Numa amostra aleatória, cada unidade tem a mesma probabilidade $\left(n/N \right)$ de ser selecionada.

Em geral, o erro padrão é inversamente proporcional a $\sqrt{n}$. Desta forma, aumentar o tamanho da amostra por um fator "*f*" reduz o erro padrão (ou seja, melhora a precisão da estimativa) por um fator de $\sqrt{f}$.

3.1 Determinação do tamanho da amostra

Pode-se minimizar o erro de amostragem aumentando-se a fração amostrada da população $\left(n/N \right)$. No entanto, aumentar a amostra causa aumento de custo, em geral. Muitas questões devem ser consideradas para otimizar o tamanho da amostra. Os seguintes passos devem ser seguidos para determinar o tamanho da amostra.

1. **Selecionar populações homogêneas**. Para que uma amostragem seja rentável, o perfil energético das unidades medidas deve ser o mesmo em toda a população. Se há dois tipos diferentes de unidades em termos de desempenho energético, deve-se formar dois grupos e amostrá-los separadamente. Por exemplo, para amostrar os períodos de operação de ambientes controlados por sensores de ocupação, os ambientes ocupados de forma mais contínua (por exemplo, escritórios com várias pessoas) devem ser separados daqueles com ocupação eventual (por exemplo, salas de reunião).
2. **Determinar os níveis desejados de precisão e confiança** para a estimativa (por exemplo, horas de uso) a determinar. A precisão refere-se ao limite de erro em torno da estimativa real (ou seja, ±x% em torno da estimativa). Melhores precisões requerem maiores amostras. A confiança refere-se à probabilidade que a estimativa fique dentro da faixa de precisão (ou seja, a probabilidade que a estimativa fique dentro da faixa de ±x% definida na declaração da precisão). Probabilidades mais altas também requerem maiores amostras. Por exemplo, se se deseja ±10% de precisão a 90% de confiabilidade (chamado critério 90/10 de incerteza), quer-se dizer que a faixa definida por ±10% em torno da média conterá o valor verdadeiro (que não se conhece) com a probabilidade de 90%, ou seja, em 90% dos casos. Como exemplo, ao estimar o tempo de uso da iluminação em uma instalação, decidiu-se usar amostragem porque seria muito caro medir as horas de uso de todos os circuitos. Medir uma amostra dos circuitos fornece uma estimativa das horas reais de utilização. Para atender ao critério 90/10 de incerteza, a faixa de estimativa das horas de uso (±10%) deverá ter 90% de chance de conter o valor real. A abordagem convencional é projetar a amostra para atingir, com um nível de confiança de 90%, ±10% de precisão. No entanto, o Plano de M&V deve levar em conta as restrições orçamentárias. Melhorar a precisão de 20% para 10%, por exemplo, aumentará em quatro vezes o custo, e para 2% em cem vezes o custo (já que o erro de amostragem é inversamente proporcional a $\sqrt{n}$). Selecionar o critério apropriado de incerteza significa balancear precisão e custo.
3. **Decidir o nível de desagregação.** Estabelecer como desagregar a população e se o critério de níveis de confiança e precisão devem ser aplicados a todas as sub-amostras ou não.
4. **Calcular o tamanho da amostra inicial**. Baseado na informação acima, uma estimativa inicial do tamanho da amostra pode ser determinado pela equação abaixo.

$$n_0 = \frac{z^2 * cv^2}{e^2} \qquad (16)$$

onde:

- n_o é a estimativa inicial do tamanho da amostra requerida, antes da medição
- *cv* é o coeficiente de variância, definido como o desvio padrão das leituras dividido pela média. Até que a média e desvio padrão da população possam ser estimados pelas amostras, usar o valor de projeto semelhante anterior, ou 0,5 se não houver dados
- *e* é o nível desejado de precisão
- *z* é o valor da distribuição normal padrão (usar a Tabela 1 com um número infinito de leituras e o nível desejado de confiança). Por exemplo, z é 1,96 para um nível de confiança de 95% (1,64 para 90%, 1,28 para 80% e 0,67 para 50% de confiança).

  NOTA: quando $n < 30$ usar a estatística *t*, quando $n \geq 30$ usar a estatística *z* (distribuição normal padrão) – elas tendem ao mesmo valor quando a amostra aumenta.

Por exemplo, para 90% de confiança com 10% de precisão e um cv de 0,5, a amostra inicial requerida é de

$$n_0 = \frac{1{,}64^2 \times 0{,}5^2}{0{,}1^2} = 67$$

Em alguns casos (por exemplo, medição de horas de uso de iluminação), pode ser interessante inicialmente medir-se uma amostra inicial pequena com o único objetivo de estimar o *cv*, para apoiar o planejamento da programação da amostra. Também valores de projetos de M&V anteriores podem ser usados como estimativas iniciais apropriadas do *cv*.

5. **Ajustar a estimativa inicial da amostra para pequenas populações**. O tamanho necessário da amostra pode ser reduzido se a população for menor que 20 vezes o tamanho da amostra. Para o exemplo acima (n_0 = 67), se a população (*N*) for somente de 200 unidades (ou seja, aproximadamente 3 vezes a amostra), o "ajuste para populações finitas" deve ser aplicado, como segue:

$$n = \frac{n_0 N}{n_0 + N} \qquad (17)$$

Aplicar este ajuste para populações finitas para o exemplo acima reduz o tamanho requerido da amostra *n*, para atender ao critério de incerteza 90/10, para 50 unidades.

6. **Finalizar o tamanho da amostra**. Como o tamanho inicial da amostra (n_o) foi determinado com um *cv* presumido, deve-se lembrar que o *cv* real da população amostrada pode ser diferente, levando a uma precisão diversa do que se esperava. Portanto, pode acontecer de ter que se medir mais unidades para atingir o critério desejado. Se o *cv* medido for menor que o *cv* considerado no passo 4, o tamanho da amostra será desnecessariamente grande. Ao contrário, se se considerou um cv menor que o real, a meta de precisão não será atingida a menos que se façam mais medições.

À medida que a medição for feita, a média e o desvio padrão das leituras deve ser calculado. O *cv* real e a amostra requerida (equações 16 e 17) devem ser recalculados. Este processo pode reduzir o processo de amostragem. Pode também levar a programar mais medições que o inicialmente previsto. Para manter o orçamento de M&V sob controle, pode ser apropriado definir um tamanho máximo de amostra. Se este máximo for atingido, o relatório de M&V deve citar o fato e a precisão real atingida.

4 Medição

As variáveis relativas à energia e variáveis independentes são frequentemente medidas como parte de um programa de M&V. Nenhum medidor é 100% preciso, por mais sofisticado que seja. A precisão dos medidores é publicada pelos fabricantes, a partir de testes em laboratórios. Dimensionar adequadamente o medidor para a faixa de medições possível garante que os dados coletados ficarão dentro de limites de erros (ou precisão) conhecidos e aceitáveis.

Os fabricantes tipicamente estabelecem a taxa de precisão como uma fração (%) da leitura atual ou como uma fração (%) do fundo da escala. Neste último caso, é importante considerar onde as leituras típicas irão cair na escala do medidor antes de calcular a precisão de uma leitura típica. Sobredimensionamento de medidores cuja precisão é estabelecida relativa ao fundo da escala irá reduzir a precisão da leitura real.

As leituras de muitos medidores irão 'derivar' com o tempo devido ao desgaste mecânico. Uma recalibração periódica contra um padrão conhecido é requerida devido a este problema. É importante manter a precisão de medidores usados no campo por meio de manutenções rotineiras, e calibração contra padrões conhecidos.

Adicionalmente à precisão do medidor em si, outros possíveis efeitos desconhecidos podem reduzir a precisão do sistema de medição:

- Má instalação do medidor, de modo que não consiga uma medição representativa da quantidade mensurada (por exemplo, a leitura de medidores de vazão é afetada pela proximidade de joelhos na tubulação)
- Erros de dados de telemetria que aleatoria ou sistematicamente 'cortam' os dados de medição

  Como resultado de tais erros de medição não quantificáveis, é importante ter em conta que a precisão informada pelo fabricante provavelmente é maior que a precisão das leituras reais no campo. Entretanto, não há modo de quantificar estes outros efeitos.

  A declaração da precisão pelo fabricante deve estar de acordo com os padrões relevantes da indústria para o seu produto. Deve-se tomar cuidado para determinar o nível de confiança usado para estabelecer a precisão do medidor. Se não informado, provavelmente será 95%.

  Quando uma medição única é utilizada no cálculo da economia, em vez da média de várias medições, componentes independentes são combinadas para determinar a incerteza. O desvio padrão do valor medido é:

$$EP = \frac{\text{Pr}\,ecisão\ relativa\ do\ medidor \times Valor\ medido}{t} \qquad (18)$$

  Onde o valor de t é baseado na maior amostra feita pelo fabricante do medidor para desenvolver sua declaração de precisão. Assim, deve-se usar o valor de t da Tabela 1 com "tamanho infinito".

  Quando se realiza uma medição com várias leituras do medidor, os valores observados contêm erros do medidor e variações do fenômeno sob medição. A média das leituras contém igualmente ambos os efeitos. O erro padrão do valor médio estimado da medição é encontrado usando-se a equação (4)

5 Combinação de componentes da incerteza

Tanto os termos com medições quanto o de ajustes na equação da economia

> Economia = (Consumo ou demanda da linha de base – Consumo ou demanda do período de determinação da economia) ± Ajustes

podem introduzir incertezas na determinação da economia. As incertezas dos componentes individuais podem ser combinadas para permitir uma declaração geral da incerteza da economia. Esta combinação pode ser calculada expressando-se a incerteza de cada componente por meio do seu erro padrão.

Os componentes devem ser independentes para que se possa usar os métodos a seguir para combinar incertezas. Independência significa que quaisquer erros aleatórios que afetam um dos componentes não estão relacionados aos erros que afetam os outros componentes.

Se a economia é a soma ou diferença de vários componentes independentes (*C*) (i.e., $Economia = C_1 \pm C_2 \pm ... \pm C_p$), o erro padrão da economia pode ser estimado por:

$$\text{EP}(Economia) = \sqrt{EP(C_1)^2 + EP(C_2)^2 + ... + EP(C_p)^2} \quad ....(19)$$

Por exemplo, se a economia é calculada pela equação:

> Economia = (Energia da linha de base ajustada – Energia do período de determinação da economia) ± Ajustes não de rotina

Como a diferença entre a energia da linha de base ajustada e a energia medida do período de determinação da economia, o erro padrão da diferença (economia) é calculado como:

EP(*Economia*) = $\sqrt{EP(\text{linha de base ajustada})^2 + EP(\text{energia determinação})^2}$

O *EP* (linha de base ajustada) vem do erro padrão da estimativa derivado da equação (11) – modelo de regressão. O *EP* (energia do período de determinação da economia) vem da precisão do medidor usando-se a equação (18).

Se a estimativa da economia é um produto de vários componentes independentes (C_i) (i.e., $Economia = C_1 * C_2 * ... * C_p$), o erro padrão relativo da economia é dado aproximadamente por:

$$\frac{EP(Economia)}{Economia} \approx \sqrt{\left(\frac{EP(C_1)}{C_1}\right)^2 + \left(\frac{EP(C_2)}{C_2}\right)^2 + ... + \left(\frac{EP(C_p)}{C_p}\right)^2} \qquad (20)$$

Um bom exemplo desta situação é a determinação da economia em AEE de iluminação com diminuição apenas da potência:

Economia = Δ Watts x Horas

Se o Plano de M&V requer a medição das horas de uso, "Horas" será um valor com erro padrão de medição. Se o Plano de M&V inclui também a medição da redução da potência, "Δ Watts" será também um valor com erro padrão. O erro padrão relativo da economia será calculado usando a fórmula acima como segue:

$$\frac{EP(Economia)}{Economia} = \sqrt{\left(\frac{EP(\Delta Watts)}{\Delta Watts}\right)^2 + \left(\frac{EP(Horas)}{Horas}\right)^2}$$

Quando vários resultados de economias são totalizados (por exemplo, economias mensais para calcular a economia anual) e todos têm o mesmo erro padrão, as equações (5) ou (19) podem ser usadas para calcular o erro padrão estimado do total.

EP Total (*Economia*) = $\sqrt{EP(economia_1)^2 + EP(economia_2)^2 + ... + EP(economia_N)^2}$ =

$$\sqrt{N} \times EP(Economia)$$

Onde *N* é o número de resultados de economias com o mesmo erro padrão somados.

Uma vez que o erro padrão da economia é determinado da forma anterior, é possível se fazer afirmações apropriadas e conclusivas sobre a incerteza inerente na economia, usando a matemática da curva de distribuição normal (Figura 1) ou os dados da Tabela 1. Por exemplo, podem-se calcular três valores:

1. Precisão absoluta ou relativa da economia total, para um dado nível de confiança (por exemplo, 90%), calculado usando-se os valores relevantes *t* da Tabela 1 e equações (7) ou (9), respectivamente.
2. Erro provável (*EPr*), definido como aquele com 50% de confiança. O erro provável representa o valor mais provável do erro. Em outras palavras, é igualmente provável que o erro verdadeiro seja maior ou menor que o *EPr* (ASHRAE, 1997). A Tabela 1 mostra que o nível de confiança de 50% é calculado com *t* = 0,67 para tamanhos de amostras maiores que 120, ou faixa de ±0,67 erros padrão do valor médio. Então a faixa do erro provável na economia, usando-se a equação (8), é ±0,67 x *EP* (Economia).
3. Limite de Confiança (*LC*) a 90%, definido como a faixa onde se está 90% certo que os efeitos aleatórios não produziram a diferença observada (e sim a AEE). Da Tabela 1, usando a equação (8), o LC é ±1,64 x *EP* (Economia) para amostras maiores que 120.

5.1 Avaliação de interações de múltiplos componentes da incerteza

As equações (19) e (20) para combinar componentes da incerteza podem ser usadas para estimar como os erros em um componente afetarão a precisão da economia total. Assim, os recursos de M&V podem ser usados de forma rentável para reduzir o erro na economia. Tais considerações no projeto da M&V devem levar em conta os custos e os efeitos na precisão da economia de possíveis melhoras na precisão de cada componente.

Softwares usados em ferramentas comuns para planilhas de dados permitem uma avaliação fácil do erro líquido associado com a combinação de múltiplos componentes da incerteza, usando-se o método de Monte Carlo. O método de Monte Carlo permite a avaliação de múltiplos cenários "e se" revelando a faixa de possíveis resultados, a sua probabilidade de ocorrência, e qual componente tem o maior efeito no resultado final. Tal análise identifica onde os recursos precisam ser alocados para controlar o erro. Um exemplo simples da análise "e se" é apresentada abaixo para uma AEE de iluminação. Uma luminária de potência nominal 96 W é trocada por outra de potência nominal 64 W. Se a luminária fica acesa 10 horas todos os dias, a economia anual é:

$$\textit{Economia anual} = \frac{(96-64)\times 10\times 365}{1.000} = 117\,kWh$$

A potência da nova luminária foi medida facilmente com precisão, comprovando os 64 W. No entanto, há muita variação entre as potências das luminárias antigas e entre as horas de uso em diferentes ambientes. A potência das luminárias antigas e as horas de uso não são facilmente medidas com boa certeza (ou baixa incerteza). Assim, a economia também não será conhecida com baixa incerteza. O desafio do projeto de M&V é determinar o impacto na economia se uma dessas quantidades estiver errada por quantidades plausíveis.

A Figura 2 mostra um análise de sensibilidade na economia para os dois parâmetros, potência da luminária antiga e horas de uso. Variou-se cada uma até ±30% e o impacto na economia é mostrado. Pode-se ver que a economia é significativamente mais sensível à variação na potência da luminária antiga que às horas de uso. Um erro de 30% na potência produz um erro de 90% na economia, enquanto que um erro de 30% nas horas de operação produzem apenas 30% de erro na economia.

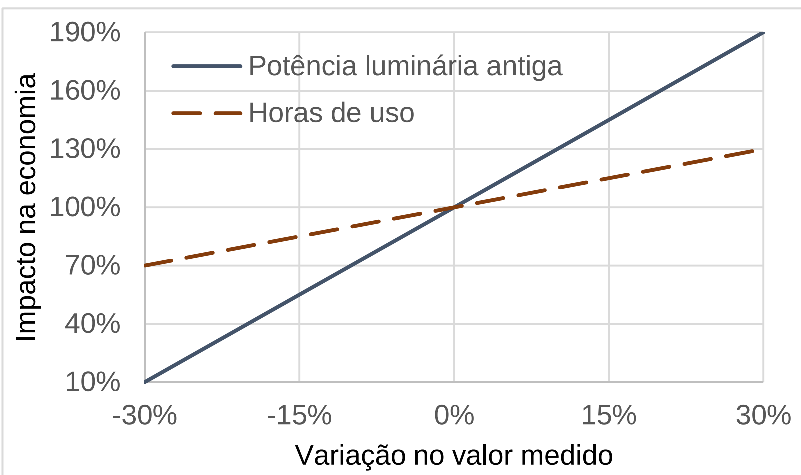

Figura 2   Exemplo de análise de sensibilidade – economia em iluminação

Se o método proposto da M&V resulta em leituras da potência da luminária antiga com uma faixa de incerteza de ±5%, a faixa na incerteza da eletricidade será de ±15%. Em outras palavras, se a potência estiver entre 91 e 100 W, a economia estará entre 99 e 135 kWh anuais. A faixa de incerteza na economia é 36 kWh (135 – 99). Se o valor marginal da eletricidade é 10 centavos por kWh, a faixa na incerteza será $3,60 anualmente. Se a precisão da luminária antiga puder ser estimada com maior precisão por um dispêndio significativamente menor que $3,60, será interessante aumentar os esforços em medição, dependendo também do número de anos em que a economia será considerada.

A Figura 2 mostra que o termo de horas de uso tem um impacto bem menor na economia final neste exemplo (a linha relativa às horas de uso é menos inclinada, indicando menor sensibilidade). É plausível que o erro na medição das horas de uso seja de ±20%, resultando numa faixa de incerteza de também ±20% ou 23 kWh anuais (20% de 117 kWh). A faixa de variação da economia é cerca de 46 kWh (= 2 x 23 kWh), o que representa $4,60 ao ano. Novamente, pode ser interessante aumentar a precisão na medição das horas de uso se isto puder ser feito por bem menos que $4,60 anuais, dependendo também do número de anos do período de determinação da economia.

A faixa dos erros possíveis na economia resultante de erros na medição nas horas de uso (46 kWh) é maior que os erros resultantes da medição da potência das velhas luminárias (36 kWh). Este é o efeito oposto que seria esperado baseado na maior sensibilidade da economia à potência do que às horas de uso, como visto na Figura 2. Essa diferença surge porque o erro plausível das horas de uso (±20%) é muito maior que o erro plausível da medição da potência da luminária antiga (±5%).

A análise de sensibilidade como a acima pode tomar várias formas. O exemplo precedente mostrou os princípios. A simulação de Monte Carlo permite considerações complexas de muitos diferentes parâmetros, permitindo ao projeto de M&V direcionar gastos onde for mais necessário para melhorar a precisão geral da economia.

5.2 Estabelecimento de metas para as incertezas quantificáveis da economia

Como discutido anteriormente, nem todas as incertezas podem ser quantificadas. Entretanto, aquelas que o podem fornecem orientação para o planejamento da M&V. Considerando o custo de M&V das várias abordagens possíveis, o programa de M&V pode produzir o tipo de informação aceitável por todos os interessados na determinação da economia, inclusive aqueles que devem pagar pela determinação. Em última análise, o Plano de M&V deve mostrar o nível esperado das incertezas quantificáveis.

A determinação da economia de energia requer estimar a diferença entre níveis de energia, mais do que simplesmente medir o nível de energia em si. Em geral, calcular a diferença para atender a um critério de precisão relativa requer uma precisão absoluta na medição dos componentes maior que a precisão absoluta requerida para a diferença. Por exemplo, suponha que a carga média está em torno de 500 kW, e a economia prevista em 100 kW. O critério de erro de ±10% com 90% de confiança ("90/10") pode ser aplicado de duas formas:

- Se aplicado à medição da carga, a precisão absoluta deve ser 50 kW (10% de 500 kW) com 90% de confiança.
- Se aplicado à economia, a precisão absoluta deve ser 10 kW (10% de 100 kW) nos mesmos 90% de confiança. Para chegar nos 10 kW de precisão na determinação da economia deve-se ter precisões absolutas de 7 kW nos componentes (usando a equação (19), se ambos componentes têm a mesma precisão).

Claramente, a aplicação do critério de confiança/precisão 90/10 na apuração da economia requer muito maior precisão na medição das cargas que o critério 90/10 aplicado à carga em si.

O critério de precisão pode ser aplicado não somente à economia de energia, porém também aos parâmetros que a determinam. Por exemplo, suponha que a economia é o produto do número de unidades (*N*), horas de operação (*H*) e redução em potência (*R*, em watts): Economia = *N* x *H* x *R*. O critério 90/10 pode ser aplicado separadamente a cada um desses parâmetros. No entanto, atingir o critério 90/10 de precisão para cada parâmetro separadamente não significa atingir o critério para a economia, que é o que interessa no fim. De fato, utilizando a equação (20) a precisão seria somente 17% a 90% de confiança. Por outro lado, se o número de unidades e redução na potência são conhecidos sem erro, a precisão de 90/10 para as horas de uso significa 90/10 para a economia.

O padrão de precisão pode ser atribuído em diversos níveis. A escolha do nível de desagregação afeta dramaticamente o projeto de M&V e os custos associados. Em geral, os requisitos de coleta de dados aumentam se os requisitos são atribuídos a cada componente. Se o principal objetivo é controlar a precisão da economia do projeto global, não é necessário impor o mesmo requisito de precisão a cada componente.

6 Exemplo de análise de incerteza

Para ilustrar o uso das várias ferramentas estatísticas para a análise de incerteza, a Tabela 3 mostra um exemplo de resultado de modelo de análise de regressão em planilha. É uma regressão de 12 consumos mensais medidos pela concessionária de energia elétrica de um prédio e os graus-dia de refrigeração (GDR) do mesmo período. É somente uma parte do resultado da planilha. Os valores de interesse estão destacados em itálico.

Tabela 3 – Exemplo de resultado de planilha com análise de regressão

RESUMO DOS RESULTADOS					
Estatística de regressão					
R múltiplo	0,97				
R-Quadrado	*0,93*				
R-quadrado ajustado	0,92				
Erro padrão	*367,5*				
Observações	12				
	Coeficientes	*Erro padrão*	*Estat t*	*95% inferiores*	*95% superiores*
Interseção	*5.634,15*	151,96	37,08	5.295,56	5.972,74
GDR	*7,94*	*0,68*	*11,64*	6,42	9,45

Para a linha de base dos 12 consumos mensais em kWh e GDR associados, o modelo de regressão calculado é:

Consumo mensal de eletridade [kWh] = 5.634,15 + (7,94 * GDR)

O coeficiente de determinação, R^2 (mostrado como R-Quadrado na Tabela 3), é bem alto (em 0,93), indicando que 93% da variação da energia nos 12 meses é explicada pelo modelo por meio dos dados de GDR. Este fato implica uma relação muito forte e que o modelo pode ser utizado para estimar os termos de ajuste na equação

Economia = (Consumo ou demanda da linha de base – Consumo ou demanda do período de determinação da economia) ± Ajustes

O coeficiente estimado de 7,94 kWh/GDR tem um erro padrão de 0,68. Esse *EP* leva a uma estatística *t* (mostrada como Estat t na Tabela 3) de 11,64. Esta estatística *t* é então comparada ao valor *t* crítico apropriado na Tabela 1 (*t* = 2,2 para 12 pontos de dados e confiança de 95%). Já que 11,64 excede 2,2, os GDR são uma variável independente com grande significado. A planilha também mostra que a faixa para o coeficiente a 95% de confiança é de 6,42 a 9,45, o que dá uma precisão relativa de ±19% (= (7,94 – 6,42) / 7,94). Em outras palavras, há uma confiança de 95% que cada GDR adicional aumenta o consumo mensal entre 6,42 e 9,45 kWh.

O erro padrão da estimativa usando a fórmula de regressão é 367,5. A média mensal de GDR no período de determinação é 162 (não mostrado no resultado). Para predizer que consumo haveria em condições de refrigeração médias, por exemplo, insere-se este valor de GDR no modelo de regressão:

Consumo previsto = 5.634 + (7,94 x 162) = 6.920 kWh por GDR médios mensais

Usando o valor de *t* da Tabela 1 de 2,2, para 12 pontos de dados e nível de confiança de 95%, a faixa das possíveis predições é:

Faixa de predições = 6.920 ± (2,2 x 367,5) = 6.112 a 7.729 kWh

A precisão absoluta é aproximadamente ±809 kWh (= 2,2 x 367,5) e a precisão relativa é ±12% (= 809/6.920). O valor apresentado pela planilha para o erro padrão da estimativa fornece a informação necessária para calcular a precisão relativa esperada com o uso do modelo de regressão para qualquer mês, neste caso 12%. Se o consumo do período de determinação da economia for 4.300 kWh, calcula-se a economia como:

Economia = (6.920 – 4.300) = 2.620 kWh

Já que o medidor da concessionária foi utilizado para obter o valor da eletricidade no período de determinação da economia, os valores medidos podem ser considerados 100% precisos (*EP* = 0%) porque o medidor da concessionária define o montante a pagar, independentemente de erros. O *EP* da economia será:

$$EP(economia\ mensal) = \sqrt{EP(linha\ de\ base\ ajustada)^2 + EP(período\ de\ \det er\min ação)^2}$$

$$EP = \sqrt{367{,}5^2 + 0^2} = 367{,}5$$

Usando-se um *t* de 2,2, a faixa de possíveis economias mensais é

Faixa de economias = 2.620 ± (2,2 x 367,5) = 2,620 ± 810 = 1.810 a 3.430 kWh/mês

Para determinar a precisão da economia anual total (soma das mensais), considerou-se que o erro padrão de cada economia mensal é o mesmo. Então a economia anual tem um erro padrão de:

EP (economia anual) = $\sqrt{12 \times 367{,}5^2}$ = 1.273 kWh

Já que o *t* provém do modelo da linha de base, conserva o valor de 2,2 usado acima. Assim a precisão absoluta da economia anual é 2,2 x 1.273 = 2.801 kWh. Se as economias mensais forem todas de 2.620 kWh, a economia anual será de 31.440 kWh, e a precisão relativa da economia anual será 9% (= 2.801 / 31.440).

A EVO gostaria de agradecer a seus subscritores institucionais:

Principais

BC Hydro
Services Industriels de Genève

Senior

EDF Electricité de France
Pacific Gas & Electric
Southern California Edison
Schneider – Electric

Governamentais

Ontario Power Authority
Bonneville Power Administration

Governamentais

ADENE – Agência para a Energia (Portugal)

Educacional

Université de Genève

Associados Anexo 1

Navigant Consulting Inc.

Quantum Energy Services &Technologies, Inc. (QuEST)

EEVS – Energy Efficiency Verification Specialists

HEP-ESCO d.o.o.

Sem fins lucrativos não Anexo 1

Taiwan Green Productivity Foundation (TGPF)