EFFICIENCY VALUATION ORGANIZATION

International Performance Measurement and Verification Protocol

(Protocollo internazionale di misura e verifica delle prestazioni)

Statistica e incertezza per IPMVP

Preparato da Efficiency Valuation Organization
www.evo-world.org

Giugno 2014

EVO 10000 – 1:2014 (IT)

Visione EVO

Un mercato globale che valorizzi correttamente l'uso efficiente delle risorse naturali e utilizzi l'efficienza negli usi finali come valida alternativa alla fornitura di energia.

Missione di EVO

Sviluppare e promuovere l'uso di protocolli, metodi e strumenti standardizzati per quantificare e gestire i rischi e i benefici delle prestazioni legate a transazioni commerciali riguardanti l'efficienza energetica negli usi finali, le energie rinnovabili e l'efficienza idrica.

Versione italiana

La versione italiana di Statistica e incertezza per IPMVP è stata tradotta da Daniele Forni e Dario Di Santo (FIRE - Federazione italiana per l'uso razionale dell'energia) e controllata da Alessandra Caccamo e Michele Liziero (Energy Team S.p.A.). EVO ringrazia tutte le persone che hanno contribuito alla traduzione.

Indice Pagina

1 Introduzione

L'obiettivo della M&V è determinare in modo attendibile i risparmi energetici. Le rendicontazioni dei risparmi richiedono un ragionevole livello d'incertezza per essere affidabili. L'incertezza di un rapporto sui risparmi può essere gestita attraverso il controllo di errori casuali e di errori sistematici dei dati. Gli errori casuali sono influenzati dalla qualità degli strumenti di misura, dalle tecniche di misura e dalla progettazione della procedura di campionamento. Gli errori sistematici dei dati sono influenzati dalla qualità dei dati di misura, dalle ipotesi e dalle analisi. Riducendo gli errori, di solito, il costo di M&V aumenta; quindi la necessità di migliorare l'incertezza deve essere giustificata dal valore di una migliore informazione.

I calcoli dei risparmi energetici comportano un confronto dei dati energetici misurati e un calcolo di "aggiustamenti" per convertire entrambe le misure allo stesso insieme di condizioni operative. Sia le misure sia gli aggiustamenti introducono errori che possono verificarsi a causa di imprecisioni del contatore, delle procedure di campionamento e delle procedure di aggiustamento. Questi processi generano "stime" statistiche con valori calcolati o attesi e un certo livello di variazione. In altre parole, non sono noti i valori veri, ma solo stime con un certo livello d'incertezza. Tutte le misure fisiche e le analisi statistiche si basano sulla stima delle tendenze centrali, come i valori medi e la quantificazione delle variazioni come l'intervallo, la deviazione standard, l'errore standard e la varianza.

La statistica è quel corpus di metodi matematici applicabili ai dati per aiutare a prendere decisioni in caso d'incertezza. La statistica fornisce modalità di verifica dei risultati allo scopo di valutare se i risparmi rendicontati sono "significativi", ovvero sono verosimilmente l'effetto di un'azione di miglioramento dell'efficienza energetica (AMEE) piuttosto che un comportamento casuale.

Si verificano tre tipi di errori: modellazione, campionamento e misura.

- **Modellazione**. Errori di modellazione matematica dovuti a forme funzionali inadeguate, inclusione di variabili non pertinenti, esclusione di variabili pertinenti, ecc.
- **Campionamento**. L'errore di campionamento si verifica quando si misura solo una porzione della popolazione dei valori reali o quando si usa un approccio di campionamento viziato da un errore sistematico. La rappresentazione di una sola porzione della popolazione può verificarsi in senso fisico o in senso temporale.
- **Misura**. Gli errori di misura sono generati dall'accuratezza dei sensori, da errori di tracciamento dei dati, da deriva post calibrazione, da misure imprecise, ecc. L'entità di tali errori è in gran parte ascrivibile alle specifiche del produttore ed è gestita con una ricalibrazione periodica.

Il presente documento fornisce indicazioni per quantificare le incertezze causate da queste tre forme di errore. Alcune fonti di errore non sono note né quantificabili:

- ➢ scelta o posizionamento inadeguati del contatore;
- ➢ stime imprecise nell'Opzione A, o
- ➢ errata stima degli effetti interattivi nelle Opzioni A o B.

Le incertezze non note o non quantificabili possono essere gestite solo seguendo le migliori pratiche del settore.

1.1 Come esprimere l'incertezza

Per essere comunicati in modo statisticamente valido, i risparmi devono essere espressi con associati i loro livelli di confidenza e precisione. La confidenza si riferisce alla possibilità o alla probabilità che i risparmi stimati rientrino nell'intervallo di precisione.[1]

[1] I termini statistici selezionati sono definiti alla sezione 1.3

ESEMPIO: Il processo di stima dei risparmi può portare a un'affermazione del tipo: "La migliore stima di risparmio è di 1.000 kWh all'anno (stima puntuale) con una probabilità del 90% (confidenza) che il valore medio del risparmio reale rientri nel ± 20% di 1000." Una presentazione grafica di questo rapporto è illustrata in Figura 1.

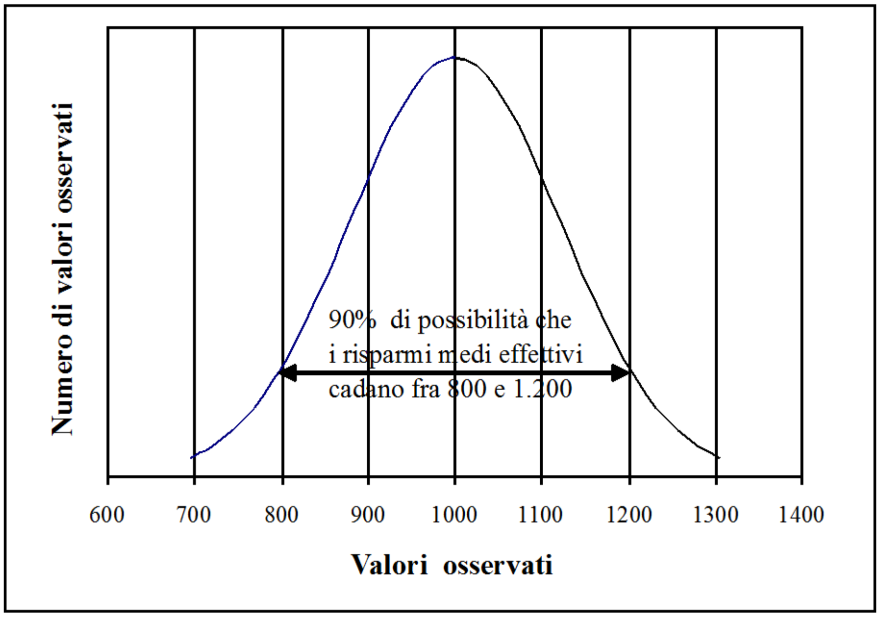

Figura 1 Distribuzione normale della popolazione

Un'indicazione della precisione statistica (porzione ± 20%) senza un livello di confidenza (porzione del 90%) è imprecisa. Il processo di M&V può generare un altissimo grado di precisione con un basso grado di confidenza.

Esempio: i risparmi si possono indicare con una precisione di ± 1%, ma il livello di confidenza associato può scendere dal 95% al 35%.

1.2 Incertezza accettabile

I risparmi sono considerati statisticamente validi se sono grandi rispetto alle variazioni statistiche. In particolare, i risparmi devono essere superiori di due volte all'errore standard del valore di riferimento. Se la varianza dei dati di riferimento è eccessiva, il comportamento casuale non spiegato del consumo di energia della struttura/impianto o del sistema è elevato e ogni singola determinazione di risparmio è inaffidabile.

Ove non fosse possibile soddisfare questo criterio, prendere in considerazione quanto segue:

- apparecchiature di misura più precise,
- un numero maggiore di variabili indipendenti nel modello matematico,
- campioni di dimensioni maggiori o
- un'Opzione IPMVP che sia meno influenzata da variabili non note.

1.3 Definizioni di termini statistici

Media del campione ($\bar{Y}$): determinata dividendo la somma dei singoli dati (Y_i) per il numero totale dei dati (n), come segue:

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

Varianza del campione (S^2): la varianza del campione indica la misura in cui i valori osservati differiscono l'uno dall'altro, cioè, la variabilità o la dispersione. Maggiore è la variabilità, maggiore l'incertezza nella media. La varianza del campione si trova facendo la media dei quadrati delle singole deviazioni dalla media. La ragione per cui queste deviazioni dalla media sono elevate al quadrato è semplicemente quella di eliminare i valori negativi (quando un valore è al di sotto della media), in modo che questi ultimi non cancellino i valori positivi (quando un valore è al di sopra della media). La varianza campione è calcolata come segue:

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

Deviazione standard del campione (s): è semplicemente la radice quadrata della varianza del campione. Ciò porta la misura della variabilità di nuovo alle unità dei dati; ad esempio, se le unità di varianza sono $(kWh)^2$, le unità di deviazione standard sarebbero kWh.

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

Errore standard del campione (SE): è la deviazione standard del campione divisa per $\sqrt{n}$. Questa misura viene usata nella stima della precisione della media del campione. Nella stragrande maggioranza dei testi di statistica, viene indicata anche come $\bar{s}$ o "deviazione standard della media del campione".

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

Deviazione standard del campione totale (s_{tot}): spesso interessano più le proprietà statistiche del totale piuttosto che la media. La deviazione standard del campione totale viene utilizzata per definire la precisione del totale del campione. È definita come radice quadrata della dimensione del campione, $\sqrt{n}$ per la deviazione standard del campione:

$$S_{tot} = \sqrt{n} \bullet S \qquad (5)$$

Coefficiente di variazione (cv): il coefficiente di variazione è semplicemente la deviazione standard di una distribuzione espressa come percentuale della media. Ad esempio, il cv del totale del campione sarebbe [s_{tot}] ÷ [totale del campione]; il cv della media del campione sarebbe [$SE\bar{Y}$] ÷ [media del campione]; ecc. La formula generale è:

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

Precisione: la precisione è la misura dell'intervallo assoluto o relativo all'interno del quale dovrebbe trovarsi il valore reale con il livello di confidenza specificato. Il livello di confidenza si riferisce alla probabilità che l'intervallo citato contenga il parametro stimato.

Precisione assoluta: viene calcolata dall'errore standard del campione con un valore di "t" dalla tabella della "distribuzione t di Student". Una tabella della distribuzione t di Student viene riportata qui di seguito, ma è reperibile anche nelle tabelle statistiche, nei libri o nelle risorse on-line.

$$t \bullet SE_{\bar{Y}} \quad (7)$$

Tabella 1 Tabella t di Student

Gradi di libertà GdL	Livello di confidenza				Gradi di libertà GdL	Livello di confidenza			
	95%	90%	80%	50%		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
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

Note: Calcolare i gradi di libertà (GdL) usando:

- GdL= n – 1 (per una distribuzione campione)
- GdL = n - p – 1 (per un modello di regressione)

Dove
n = dimensioni del campione
p = numero di variabili del modello di regressione

In generale, il valore reale di qualsiasi stima statistica è previsto con un determinato livello di confidenza che rientra nell'intervallo definita da

Intervallo = stima ± *precisione assoluta* (8)

Dove "stima" è un valore derivato empiricamente di un parametro d'interesse (ad esempio consumo totale, numero medio di unità prodotte, ecc.).

Precisione relativa è la precisione assoluta diviso la stima:

$$\frac{t \bullet SE}{stima} \quad (9)$$

Esempio: considerare i dati di cui alla Tabella 2 desunti dalle 12 letture mensili di un contatore e la relativa analisi delle differenze tra ogni lettura

Tabella 2 Esempio di analisi di dati

	Lettura effettiva	Differenze calcolate rispetto alla media	
		Differenza semplice	Differenza al quadrato
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
Totale	12.000		246.600

Il Valore medio è: $\bar{Y} = \frac{\sum Y_i}{n} = \frac{12.000}{12} = 1.000$

La Varianza è: $S^2 = \frac{\sum (Y_i - \bar{Y})^2}{n-1} = \frac{246.600}{12-1} = 22.418$

La Deviazione standard è: $s = \sqrt{S^2} = \sqrt{22.418} = 150$

L'Errore standard è: $SE = \frac{s}{\sqrt{n}} = \frac{150}{\sqrt{12}} = 43$

Nella Tabella 2, ci sono 12 dati, il che significa GdL=12-1=11. Utilizzando la Tabella 1, per un livello di confidenza del 90%, il valore di "t" è 1,80. Pertanto:

la Precisione assoluta è: $t \bullet SE = 1{,}80 \times 43 = 77$

la Precisione relativa è: $\frac{t \bullet SE}{stima} = \frac{77}{1.000} = 7{,}7\%$

Quindi c'è un livello di confidenza del 90% del fatto che il consumo mensile medio effettivo si trovi nell'intervallo tra 923 e 1.077 kWh. Si può dire con un livello di confidenza del 90%, che il valore medio delle 12 osservazioni è 1.000 ± 7,7%. Analogamente si potrebbe dire che:

- con una confidenza del 95%, il valore medio delle 12 osservazioni è 1.000 ±9,5%, o
- con una confidenza dell'80%, il valore medio delle 12 osservazioni è 1.000 ±5,8%, o
- con una confidenza del 50%, il valore medio delle 12 osservazioni è 1.000 ±3,0%.

2 Modellazione

La modellazione matematica viene usata nella M&V per preparare il termine degli aggiustamenti ordinari nelle varie versioni per i risparmi trattati nei Concetti base dell'o PMVP e così riassunti:

1. Risparmi = (Consumo o domanda del periodo di riferimento – Consumo o domanda del periodo di rendicontazione) ± Aggiustamenti
2. Risparmi = (Energia del periodo di riferimento – Energia periodo di rendicontazione) ± Aggiustamenti ordinari ± Aggiustamenti straordinari
3. Consumo energetico evitato (o Risparmi) = (Energia del periodo di riferimento ± Aggiustamenti ordinari alle condizioni del periodo di rendicontazione ± Aggiustamenti straordinari alle condizioni del periodo di rendicontazione) – Energia del periodo di rendicontazione
4. Consumo energetico evitato (o Risparmi) = (Energia del periodo di riferimento adeguata – Energia periodo di rendicontazione ± Aggiustamenti straordinari dell'energia del periodo di riferimento alle condizioni del periodo di rendicontazione
5. Risparmi normalizzati = (Energia del periodo di riferimento ± Aggiustamenti ordinari alle condizioni fissate ± Aggiustamenti straordinari alle condizioni fissate) – (Energia periodo di rendicontazione ± Aggiustamenti ordinari alle condizioni fissate ± Aggiustamenti straordinari alle condizioni fissate)
6. Risparmi Opzione A = Valore stimato x (Parametro misurato nel periodo di riferimento, – Parametro misurato nel periodo di rendicontazione)
7. Risparmi Opzione B = Energia del periodo di riferimento – Energia del periodo di rendicontazione
8. Risparmi = Energia di riferimento dal modello calibrato [valore ipotetico o senza AMEE] – Energia del periodo di rendicontazione dal modello calibrato [con AMEE]
9. Risparmi = Energia di riferimento dal modello calibrato [valore ipotetico o senza AMEE] – Energia effettiva del periodo di calibrazione ± Errore di calibrazione nella lettura di calibrazione corrispondente

La modellazione implica l'individuazione di una relazione matematica tra variabili dipendenti e variabili indipendenti. La variabile dipendente, di solito l'energia, è modellata come se governata da una o più variabili indipendenti (s) X_i (note anche come variabili 'esplicative'). Questo tipo di modellazione è denominato analisi di regressione.

Nell'analisi di regressione, il modello cerca di "spiegare" la variazione di energia derivante da variazioni nelle singole variabili indipendenti (X_i). Ad esempio, se una delle X è il livello di produzione, il modello valuterebbe se la variazione di energia rispetto alla sua media è causata da variazioni del livello di produzione. Il modello quantifica la causalità. Ad esempio, quando la produzione aumenta di una unità, il consumo di energia aumenta di "b" unità, dove "b" è denominato coefficiente di regressione.

I modelli più comuni sono le regressioni lineari della forma:

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

dove:

- Y è la variabile dipendente, di solito in forma di consumo energetico durante uno specifico periodo di tempo (ad esempio 30 giorni, 1 settimana, 1 giorno, 1 ora, ecc.)
- X_i (i = 1, 2, 3, … p) rappresenta le variabili indipendenti 'p' come meteo, produzione, occupazione, durata del periodo di misura, ecc.
- b_i (i = 0, 1, 2, … p) rappresenta i coefficienti derivati per ciascuna variabile indipendente e un coefficiente fisso (b_0) non correlato alle variabili indipendenti
- e rappresenta gli errori residui che rimangono inspiegabili dopo aver considerato l'impatto delle diverse variabili indipendenti. L'analisi di regressione trova l'insieme dei valori b_i che minimizza la somma dei termini quadratici di errore residuo (quindi i modelli di regressione sono chiamati anche modelli dei minimi quadrati).

Un esempio del modello sopra riportato per il consumo energetico di un edificio è:

$$\text{consumo energetico mensile} = 342.000 + (63 \times GG_{ris}) + (103 \times GG_{raf}) + (222 \times \text{Occupazione})$$

GG_{ris} e GG_{raf} sono i gradi giorno di riscaldamento (GG_{ris}) e di raffrescamento (GG_{raf}). L'Occupazione è una misura dell'occupazione percentuale dell'edificio. In questa modalità, 342.000 è una stima del carico di base in kWh, 63 misura la variazione dei consumi per un GG_{ris} aggiuntivo, 103 misura la variazione dei consumi per un GG_{raf} aggiuntivo e 222 misura il cambiamento nel consumo per ciascuna variazione dello 1% dell'occupazione.

2.1 Errori di modellazione

Quando si usano i modelli di regressione, come sopra descritto, possono essere introdotti vari tipi d'errore, come qui di seguito elencato:

1. Il modello è costruito su valori al di fuori del probabile intervallo delle variabili da utilizzare. Dovrà essere realizzato un modello matematico che si avvalga solo di valori ragionevoli delle variabili dipendenti e indipendenti.
2. Il modello matematico può non includere variabili indipendenti pertinenti, con possibilità di introdurre relazioni basate su errori sistematici (errori sistematici dovuti alle variabili omesse).
3. Il modello può includere alcune variabili non pertinenti.
4. Il modello può avvalersi di una forma funzionale inadeguata.
5. Il modello può basarsi su dati non sufficienti o non rappresentativi.

Questi errori sono trattati più dettagliatamente nel testo sotto riportato.

2.1.1 Uso di dati fuori intervallo

Se il modello si basa su dati che non sono rappresentativi del normale comportamento energetico dell'impianto/struttura le previsioni non possono essere affidabili. Ciò può comprendere l'inclusione di valori anomali o di valori ben al di fuori dall'intervallo ragionevole. I dati dovranno essere vagliati prima di costruire il modello.

2.1.2 Omissione di variabili pertinenti

Nella M&V, l'analisi di regressione viene utilizzata per tenere conto delle variazioni d'uso energetico. La maggior parte dei sistemi complessi che utilizzano energia è influenzata da innumerevoli variabili indipendenti. I modelli di regressione non possono pretendere di includere tutte le variabili indipendenti. Anche se fosse possibile, il modello sarebbe troppo complesso per essere utile e richiederebbe un'eccessiva attività di raccolta dei dati. L'approccio pratico è includere solo una o più variabili indipendenti che si ritiene possano avere un impatto significativo sull'energia.
L'omissione di una variabile indipendente pertinente può costituire un errore importante. Se manca una variabile indipendente pertinente (ad esempio, GG_{ris}, produzione, occupazione), allora il modello non riuscirà a rappresentare una porzione significativa della variazione di energia. Il modello carente inoltre attribuirà alcune delle variazioni dovute alla variabile mancante alla variabile o alle variabili incluse nel modello, con la conseguenza di avere un modello meno accurato.

Non vi sono indicazioni evidenti di questo problema nei test statistici standard (tranne forse un R^2 basso). L'esperienza e la conoscenza delle tecnologie del sistema di cui vengono misurate le prestazioni sono elementi preziosi nell'affrontare questo problema.

Ci possono essere casi in cui è nota l'esistenza di una relazione con una variabile registrata durante il periodo di riferimento, tuttavia la variabile non è inclusa nel modello a causa della mancanza di budget per una raccolta dati continua nel periodo di rendicontazione. Tale omissione di una variabile rilevante va annotata e giustificata nel Piano di M&V.

2.1.3 Inclusione di variabili non pertinenti

A volte i modelli includono una o più variabili indipendenti non pertinenti. Se la variabile non pertinente non ha alcun rapporto (correlazione) con le variabili pertinenti incluse, allora avrà un impatto minimo sul modello. Tuttavia, se la variabile non pertinente è correlata ad altre variabili pertinenti nel modello, può influenzare i coefficienti delle variabili pertinenti.

Prestare attenzione ad aggiungere più variabili indipendenti in un'analisi di regressione solo perché sono disponibili. Giudicare la pertinenza delle variabili indipendenti richiede esperienza ed intuizione. Tuttavia, il t di Student associato è un modo per confermare la pertinenza di particolari variabili indipendenti incluse in un modello. Per determinare la pertinenza delle variabili indipendenti, occorre aver acquisito esperienza nell'analisi energetica per il tipo di impianto/struttura coinvolta in qualsiasi programma di M&V.

2.1.4 Forma funzionale

È possibile modellare una relazione usando una forma funzionale non corretta. Ad esempio, una relazione lineare potrebbe essere usata in modo errato nella modellazione di un rapporto fisico sottostante non lineare. Ad esempio in edifici riscaldati e raffrescati elettricamente su un periodo di un anno il consumo di energia elettrica e la temperatura ambiente tendono ad avere una relazione non lineare (spesso con una forma a 'U') con la temperatura esterna. (Il consumo di energia elettrica è elevato sia per temperature esterne alte che basse, mentre è relativamente basso nelle mezze stagioni.) Modellare questa relazione non lineare con un modello lineare singolo genererebbe errori inutili. Si devono invece derivare modelli lineari distinti per ogni stagione.

Può anche essere opportuno cercare relazioni d'ordine superiore, ad esempio, $Y = f(X, X^2, X^3)$.
Il modellatore deve valutare diverse forme funzionali e scegliere quella più appropriata utilizzando misure di valutazione.

2.1.5 Insufficienza dei dati

Gli errori possono verificarsi anche per dati insufficienti in termini di quantità (cioè troppi pochi punti di dati) o tempo (ad esempio, utilizzando i mesi estivi nel modello e cercando di estrapolare i mesi invernali). I dati utilizzati nella modellazione devono essere rappresentativi dell'intervallo di operazioni dell' impianto/struttura. Il periodo di tempo coperto dal modello deve includere varie stagioni possibili, tipi di utilizzo, ecc. Ciò può richiedere sia l'estensione dei periodi temporali utilizzati, sia l'aumento dimensionale del campione.

2.2 Valutazione dei modelli di regressione

Per valutare quanto validamente un particolare modello di regressione spieghi la relazione tra consumo energetico e variabile o variabili indipendenti, si possono eseguire tre test, come descritto di seguito.

2.2.1 Coefficiente di determinazione (R^2)

Il primo passo nella valutazione dell'accuratezza di un modello è esaminare il coefficiente di determinazione, R^2, misura del grado in cui le variazioni della variabile dipendente Y dal suo valore medio sono spiegate dal modello di regressione. Dal punto di vista matematico, R^2 è:

$$R^2 = \frac{\textit{Variazione spiegata di } Y}{\textit{Totale variazione di } Y}$$

o più esplicitamente:

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

dove:

- $\hat{Y}_i$ = valore energetico predetto dal modello per un particolare punto, utilizzando il valore misurato della variabile indipendente (cioè ottenuto sostituendo i valori X nel modello di regressione)
- $\overline{Y}$ = media dei valori energetici n misurati, calcolata con l'equazione (1)
- Y_i = valore energetico effettivamente osservato (cioè misurato con un contatore)

Tutti i pacchetti statistici e le funzioni di analisi di regressione dei fogli di calcolo calcolano il valore di R^2.

L'intervallo dei possibili valori per R^2 va da 0,0 a 1,0. Un R^2 di 0,0 significa che nessuna variazione è spiegata dal modello, che quindi non fornisce alcuna indicazione per comprendere le variazioni di Y (cioè, la variabile o le variabili indipendenti selezionate non forniscono spiegazioni delle cause delle variazioni osservate in Y). Al contrario, un R^2 di 1,0 significa che il modello spiega il 100% delle variazioni di Y (ossia, il modello prevede Y con certezza totale per qualsiasi insieme di valori di variabile o di variabili indipendenti). Nessuno di questi valori limite di R^2 è probabile con dati reali.

In generale, maggiore è il coefficiente di determinazione, migliore è la capacità del modello di descrivere la relazione tra le variabili indipendenti e la variabile dipendente che offre il modello. Anche se non esiste uno standard universale per un valore minimo di R^2, 0,75 è spesso considerato un indicatore ragionevole di un buon rapporto di causalità fra consumo di energia e variabili indipendenti.

Il test R^2 deve essere utilizzato solo come controllo iniziale. I modelli non devono essere respinti o accettati unicamente sulla base di R^2. Infine, un R^2 basso indica che non sono state incluse una o più variabili pertinenti o che la forma funzionale del modello (ad esempio, lineare) non è appropriata. In questa situazione sarebbe logico considerare delle variabili indipendenti supplementari o una diversa forma funzionale.

2.2.2 Errore standard della stima

Quando un modello viene utilizzato per prevedere un valore del consumo di energia (Y) per una o più variabili indipendenti, l'accuratezza della previsione è misurata dall'errore standard della stima ($SE_{\hat{Y}}$). Questa misura di accuratezza è fornita da tutti i pacchetti di regressione standard e fogli di calcolo.
Una volta che il valore o i valori della o delle variabili indipendenti sono inseriti nel modello di regressione per stimare un valore del consumo di energia ($\hat{Y}$), si può calcolare approssimativamente l'intervallo dei possibili valori per $\hat{Y}$ usando l'equazione 8 come:

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

dove:

- $\hat{Y}$ è il valore predetto del consumo di energia (Y) dal modello di regressione
- t è il valore ottenuto dalla tabella t di Student (vedere Tabella 1)
- $SE_{\hat{Y}}$ è l'errore standard della stima (previsione), calcolato come:

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

dove p è il numero di variabili indipendenti dell'equazione di regressione.

Questo parametro statistico viene spesso definito come la radice quadrata dell'errore quadratico medio (RMSE). Dividendo lo RMSE per il consumo energetico medio si ottiene il coefficiente di variazione dello RMSE o CV(RMSE).

$$CV(RMSE) = \frac{SE_{\hat{Y}}}{\overline{Y}} \quad (12)$$

Una misura statistica simile è l'errore sistematico medio (MBE) definito come:

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

MBE è un buon indicatore dell'errore sistematico complessivo nella stima della regressione. Un MBE positivo indica che le stime di regressione tendono a sovrastimare i valori effettivi. Un errore sistematico complessivo positivo tende a cancellare l'errore sistematico negativo. Lo RMSE non presenta questo problema di cancellazione.
Tutte e tre le misure possono essere utilizzate per valutare la calibrazione di modelli di simulazione nella Opzione D.

2.2.3 t di Student

Dato che i coefficienti di modello di regressione (b_k) sono stime statistiche della vera relazione tra una variabile X individuale e Y, essi sono soggetti a variazioni. L'accuratezza della stima è misurata dall'errore standard del coefficiente e dal valore associato del t di Student. Il t di Student è un test statistico che serve a determinare se una stima ha significatività statistica. Una volta stimato utilizzando il test, il valore potrà essere confrontato con i valori critici della tabella t di Student (Tabella 1).

L'errore standard di ciascun coefficiente è calcolato dal software di regressione. La seguente equazione vale per il caso di una variabile indipendente.

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

Per i casi con più di una variabile indipendente, l'equazione fornisce una ragionevole approssimazione quando le variabili indipendenti sono veramente indipendenti (cioè non correlate). In caso contrario, l'equazione diventa molto complessa e all'analista della M&V conviene utilizzare un pacchetto software per calcolare gli errori standard dei coefficienti. L'intervallo entro il quale cade il vero valore del coefficiente b si trova con la formula (8) come:

$$b \pm t \times SE_b$$

L'errore standard del coefficiente b porta anche al calcolo del t di Student. Questo test determina in ultima analisi se il coefficiente calcolato è statisticamente significativo. Il t di Student viene calcolato da tutti i software statistici utilizzando la seguente equazione:

$$\text{t di Student} = \frac{b}{SE_b} \quad (15)$$

Una volta stimato il t di Student, può venir confrontato con i valori critici della Tabella 1. Se il valore assoluto del t di Student supera il numero appropriato della Tabella 1, allora si dovrà concludere che la stima è statisticamente valida.

Secondo una regola empirica il valore assoluto di un risultato del t di Student di 2 o superiore implica che il coefficiente stimato è significativo rispetto al suo errore standard e quindi che esiste una relazione tra Y e il particolare X associato al coefficiente. Si può quindi concludere che il b stimato non è zero.
Tuttavia, in un t di Student di circa 2, la precisione del valore del coefficiente è di circa ± 100%, ovvero non un grande attestato di confidenza nel valore di b. Per ottenere una maggiore accuratezza, diciamo di ± 10%, i valori dei t di Student devono essere attorno a 20 o l'errore standard di b deve essere non di più di 0,1 di b stesso.

Si possono prendere in considerazione le seguenti azioni per migliorare il risultato del t di Student:

- selezionare la o le variabli indipendenti con il rapporto più rilevante con il consumo di energia;
- selezionare una o più variabili indipendenti i cui valori coprano l'intervallo più ampio possibile (se X non varia nel modello di regressione, b non può essere stimato e il t di Student sarà modesto);
- raccogliere e utilizzare più dati per sviluppare il modello; o
- selezionare una forma funzionale diversa per il modello; ad esempio una forma che determini separatamente uno o più coefficienti per ciascuna stagione in un edificio che sia significativamente influenzato dalle variazioni climatiche stagionali.

3 Campionamento

Il campionamento crea errori perché non tutte le unità oggetto dello studio sono misurate. La situazione di campionamento più semplice è selezionare casualmente n unità da una popolazione totale di N unità. In un campione casuale, ogni unità ha la stessa probabilità $\left(n/N\right)$ di essere inclusa nel campione.

In generale, l'errore standard è inversamente proporzionale a $\sqrt{n}$. Ciò significa che, aumentando la dimensione del campione di un fattore "f", si riduce l'errore standard (migliorando la precisione della stima) di un fattore di $\sqrt{f}$.

3.1 Determinazione delle dimensioni del campione

È possibile ridurre l'errore di campionamento aumentando la frazione della popolazione che viene campionata $\left(n/N\right)$.

Aumentare la dimensione del campione tipicamente aumenta i costi. Vari aspetti sono critici per ottimizzare le dimensioni del campione. Nel definire la dimensione del campione, dovrebbero essere seguiti i punti qui elencati:

1. **Selezionare una popolazione omogenea.** Per rendere economicamente conveniente un campionamento le unità misurate dovrebbero essere rappresentative dell'intera popolazione. Se ci sono due diversi tipi di unità nella popolazione, dovrebbero essere raggruppate e campionate separatamente. Ad esempio, quando si progetta un programma di campionamento per misurare i periodi di funzionamento dell'illuminazione di una stanza controllata da sensori di presenza, le stanze occupate più o meno continuativamente (ad esempio uffici con più persone) devono essere campionate separatamente da quelle occupate solo occasionalmente (ad esempio sale riunioni).
2. **Determinare i livelli auspicati di precisione e confidenza** della stima (ad esempio, ore di utilizzo) che devono venir considerati. La precisione si riferisce all'errore rispetto alla stima reale (cioè intervallo di ± %x rispetto alla stima). Una maggiore precisione richiede campioni più ampi. La confidenza si riferisce alla probabilità che la stima rientri nell'intervallo di precisione (cioè la probabilità che la stima si collochi nell'intervallo ± x% definito dalla precisione indicata). Una maggiore probabilità richiede ugualmente campioni più ampi. Ad esempio, se si desidera una confidenza del 90% e una precisione di ± 10%, significa che l'intervallo definito per la stima (± 10%) conterrà il vero valore per tutto il gruppo (che non viene osservato) con una probabilità del 90%. Come esempio, stimando le ore di illuminazione in un impianto/struttura, si è deciso di utilizzare il campionamento perché era troppo costoso misurare le ore di funzionamento di tutti i circuiti di

illuminazione. La misura di un campione di circuiti ha fornito una stima delle ore di funzionamento reali. Per soddisfare un criterio 90/10 di incertezza (confidenza e precisione) viene determinata la dimensione del campione in modo che, una volta che le ore di funzionamento siano stimate con il campionamento, l'intervallo di stima del campione (± 10%) dovrà avere una probabilità del 90% di contenere le ore di utilizzo reali. L'approccio convenzionale è progettare un campionamento per ottenere un livello di confidenza del 90% e ± 10% di precisione. Tuttavia il Piano di M&V deve prendere in considerazione i limiti imposti dal budget. Migliorare la precisione diciamo da ± 20% a ± 10% aumenta la dimensione del campione di 4 volte, mentre portarla al ± 2% aumenterà la dimensione del campione di 100 volte (l'errore di campionamento è inversamente proporzionale a $\sqrt{n}$). La selezione dei criteri di campionamento appropriati richiede il bilanciamento delle esigenze di accuratezza con i costi di M&V.

3. **Decidere il livello di disaggregazione**. Stabilire se i criteri di livello di confidenza e di precisione debbano essere applicati alla misura di tutti i componenti oppure ai vari sottogruppi di componenti.

4. **Calcolare le dimensioni iniziali del campione**. Sulla base delle informazioni di cui sopra, una prima stima delle dimensioni complessive del campione può essere determinata utilizzando la seguente equazione:

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

dove:

- n_o è la stima iniziale delle dimensioni richieste del campione prima del campionamento
- cv è il coefficiente di varianza, definito come deviazione standard delle letture diviso la media. Finché la media e la deviazione standard effettive della popolazione possono essere stimate da campioni reali, 0,5 può essere utilizzato come stima iniziale per cv.
- e è il livello di precisione desiderato.
- z è il valore della distribuzione normale standard della Tabella 1, con un numero infinito di letture e per il livello di confidenza desiderato. Ad esempio z è 1,96 per un livello di confidenza del 95% (1,64 per il 90%, 1,28 per lo 80% e 0,67 per il 50% di confidenza).

  NOTA: Se n< 30 usare t di Student; se n> 30 usare il test per la distribuzione normale standardizzata (test z). Se le dimensioni del campione sono infinite, allora t di Student e test z sono uguali.

  Ad esempio, per una confidenza del 90% con una perecisone del 10% e un cv di 0,5, la stima iniziale delle dimensioni richieste del campione (n_o) è

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

In alcuni casi (ad esempio la misura di ore o uso dei sistemi di illuminazione) può essere opportuno iniziare con un piccolo campione per il solo scopo di stimare un valore cv che serva nella pianificazione del programma di campionamento. Si possono anche usare valori desunti da precedenti lavori di M&V come opportune stime iniziali di cv.

5. **Adeguare la stima delle dimensioni iniziali del campione per piccole popolazioni.** La dimensione necessaria del campione può essere ridotta se l'intera popolazione campionata è non più di 20 volte la dimensione del campione. Per l'esempio delle dimensioni iniziali del campione qui sopra, (n_o = 67), se la popolazione (N) da cui viene campionato è solo di 200, la popolazione è solo 3 volte la dimensione del campione. Pertanto può essere applicato un "aggiustamento per popolazione finita". Questo aggiustamento riduce la dimensione del campione (n) come segue:

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

L'applicazione di questo aggiustamento per popolazione finita del campione sopra riportato riduce le dimensioni del campione (n) richieste per soddisfare il criterio 90%/±10% a 50.

6. **Finalizzare le dimensioni del campione**. Poiché le dimensioni iniziali del campione (n_o) sono determinate con cv presunto, è fondamentale ricordare che il cv reale della popolazione campionata può essere diverso. Quindi, per soddisfare il criterio di precisione, possono occorrere dimensioni effettive diverse del campione. Se il cv attuale risulta essere inferiore a quello inizialmente ipotizzato al punto 4, le dimensioni del campione richiesto saranno inutilmente grandi per raggiungere gli obiettivi di precisione. Se il cv effettivo risulta maggiore di quanto ipotizzato l'obiettivo di precisione non sarà raggiunto, a meno che le dimensioni del campione non aumentino oltre il valore calcolato dalle equazioni (16) e (17).

Proseguendo il campionamento, si dovranno calcolare la media e la deviazione standard delle letture. Il cv reale e le dimensioni richieste del campione (equazioni 16 e 17) devono essere ricalcolati. Questo ricalcolo può permettere una riduzione precoce del processo di campionamento, ma può anche portare alla necessità di effettuare più campionamenti di quanti fossero originariamente previsti. Per mantenere i costi di M&V nei limiti del budget può essere opportuno stabilire una dimensione massima del campione. Se questo massimo viene infine raggiunto dopo i ricalcoli di cui sopra, il rendiconto o i rendiconti di risparmio dovranno contenere una nota sulla precisione effettiva raggiunta dal campionamento.

4 Misura

Le quantità di energia e le variabili indipendenti sono spesso misurate nell'ambito di un programma di M&V utilizzando dei contatori. Nessun contatore è accurato al 100%, anche se contatori più sofisticati possono portare l'accuratezza verso il 100%. I valori di accuratezza dei contatori selezionati è indicata dal produttore dei contatori sulla base di prove di laboratorio. Un corretto dimensionamento del contatore per l'intervallo di possibili grandezze da misurare assicura che i dati raccolti cadano nei limiti d'errore noti e accettabili (o di precisione).

Normalmente i produttori indicano la precisione come frazione della lettura corrente o come frazione del valore massimo della scala del contatore. In quest'ultimo caso è importante considerare dove si situano le letture tipiche nella scala dello strumento prima di calcolare la precisione delle letture tipiche. Il sovradimensionamento dei contatori la cui precisione viene indicata rispetto alla lettura massima ridurrà notevolmente la precisione della misura effettiva.

Le letture di molti sistemi di contatori tendono a una 'deriva' nel tempo a causa dell'usura meccanica. Una ricalibrazione periodica secondo uno standard noto è necessaria per regolare questa deriva. È importante mantenere la precisione dei contatori sul campo attraverso interventi di manutenzione ordinaria e calibrazione secondo standard noti.

Oltre all'accuratezza del contatore stesso, altri effetti eventualmente non noti possono ridurre la precisione del sistema di misura:

- cattivo posizionamento del contatore, che impedisce una 'visione' rappresentativa della quantità che si suppone di dover misurare (ad esempio, le letture di un flussometro sono influenzate dalla vicinanza a un gomito del tubo).
- errori di telemetria dei dati che casualmente o sistematicamente alterano i dati del contatore.

A seguito di tali errori di misura non quantificabili, è importante rendersi conto che la precisione indicata dal costruttore probabilmente sovrastima la precisione dei dati reali sul campo. Tuttavia non vi è alcun modo per quantificare questi effetti.

Le dichiarazioni di precisione del costruttore dovranno essere conformi allo standard del settore pertinente al prodotto in questione. Occorre prestare attenzione alla determinazione del livello di confidenza usato nel definire il grado di precisione di un contatore. Salvo diversa indicazione è verosimile che la confidenza sia del 95%.

Se in un calcolo del risparmio viene utilizzata una singola misura piuttosto che la media di diverse misure, le componenti indipendenti sono combinate per determinare le incertezze. L'errore standard del valore misurato è:

$$SE = \frac{precisione\ relativa\ contatore \text{ x } valore\ misurato}{t} \qquad (18)$$

Dove t è basato sul grande campionamento fatto dal costruttore del contatore nel definire la sua dichiarazione di precisione relativa. Pertanto il valore t di cui alla Tabella 1 dovrà essere quello per dimensioni infinite del campione.

Se si effettuano letture multiple con un contatore, i valori osservati contengono sia l'errore strumentale sia le variazioni insite nel fenomeno da misurare. La media delle letture presenta altresì entrambi gli effetti. L'errore standard del valore medio stimato delle misure si trova usando l'equazione (4).

5 Combinazione di componenti d'incertezza

Sia i componenti di misura che d'aggiustamento dell'equazione:

Risparmi = (Consumo o domanda del periodo di riferimento – Consumo o domanda del periodo di rendicontazione) ± Aggiustamenti

possono generare incertezza nella rendicontazione dei risparmi. Le incertezze dei singoli componenti possono essere combinate per consentire indicazioni globali sull'incertezza dei risparmi. Questa combinazione può essere effettuata esprimendo l'incertezza di ciascun componente in termini del proprio errore standard.

I componenti devono essere indipendenti per utilizzare i seguenti metodi per combinare le incertezze. Indipendenza significa che qualsiasi errore casuale che influenzi uno dei componenti non è correlato agli errori che influenzano altri componenti.

Se i risparmi rendicontati sono la somma o la differenza di vari componenti determinati in modo indipendente (C) (cioè, $Risparmi = C_1 \pm C_2 \pm ... \pm C_p$) l'errore standard dei risparmi rendicontati può essere stimato da:

$$\text{SE}(Risparmi) = \sqrt{SE(C_1)^2 + SE(C_2)^2 + ... + SE(C_p)^2} \qquad (19)$$

Ad esempio, se i risparmi vengono calcolati usando l'equazione:

Risparmi = (Energia del periodo di riferimento – Energia del periodo di rendicontazione) ± Aggiustamenti ordinari ± Aggiustamenti non ordinari

come differenza fra energia del periodo di riferimento adeguata e energia del periodo di rendicontazione misurata, l'errore standard della differenza (risparmi) viene calcolato come:

$$SE(Risparmi) = \sqrt{SE(energia\ riferimento\ adeguata)^2 + SE(energia\ rendicontazione)^2}$$

Lo SE (energia del periodo di riferimento adeguata) deriva dall'errore standard della stima desunta dall'equazione (11). Lo SE (energia del periodo di rendicontazione) deriva dalla precisione del contatore utilizzando l'equazione (18).

Se la stima dei risparmi rendicontati è il prodotto di vari componenti determinati in modo indipendente (C_i) (cioè $Risparmi = C_1 * C_2 * ... * C_p$), l'errore standard relativo dei risparmi viene fornito approssimativamente da:

$$\frac{SE(Risparmi)}{Risparmi} \approx \sqrt{\left(\frac{SE(C_1)}{C_1}\right)^2 + \left(\frac{SE(C_2)}{C_2}\right)^2 + ... + \left(\frac{SE(C_p)}{C_p}\right)^2} \quad (20)$$

Un valido esempio di questa situazione è la determinazione dei *risparmi* di illuminazione come:

Risparmi = Δ watt x ore

Se il Piano di M&V richiede la misura di ore di utilizzo, "ore" sarà un valore con un errore standard. Se il Piano di M&V comprende anche la misura della variazione della potenza in watt, anche il valore Δwatt conterrà un errore standard. L'errore standard relativo dei risparmi sarà calcolato utilizzando la formula di cui sopra come segue:

$$\frac{SE(Risparmi)}{Risparmi} = \sqrt{\left(\frac{SE(\Delta watt)}{\Delta watt}\right)^2 + \left(\frac{SE(ore)}{ore}\right)^2}$$

Se una serie di risultati di risparmio viene sommata e tutti i valori hanno il medesimo errore standard, si potrà usare l'equazione (5) o (19) per trovare l'errore standard stimato del totale rendicontato.

$$\text{SE (Totale dei } Risparmi\text{)} = \sqrt{SE(risparmi_1)^2 + SE(risparmi_2)^2 + ... + SE(risparmi_N)^2} =$$

$$\sqrt{N} \times SE(Risparmi)$$

Dove N è il numero dei risultati di risparmio con lo stesso errore standard che viene sommato. Una volta determinato l'errore standard dei risparmi in base al calcolo di cui sopra, sarà possibile fornire opportune indicazioni conclusive circa la quantità relativa d'incertezza insita nei risparmi, utilizzando la matematica della curva di distribuzione normale standard, Figura 1, o i dati della Tabella 1. Ad esempio, si possono calcolare tre valori:

1. <u>Precisione assoluta o relativa</u> del risparmio totale, per un determinato livello di confidenza (ad esempio, del 90%), calcolata usando il pertinente valore t della Tabella 1 e rispettivamente l'equazione (7) o (9).
2. <u>Probabile Errore (PE)</u>, definito come intervallo di confidenza del 50%, rappresenta la più probabile quantità di errore. Cioè è altrettanto probabile che l'errore sia maggiore o minore del PE (ASHRAE, 1997). La Tabella 1 mostra che il livello di confidenza del 50% viene raggiunto a t = 0,67 per dimensioni di campione superiori a 120, o errori standard di 0,67 rispetto al valore medio. Quindi l'intervallo di probabile errore nei risparmi rendicontati utilizzando l'equazione (8) è di ± 0,67 x SE (Risparmi).
3. <u>Limite di confidenza del 90%</u> (CL), definito come intervallo in cui si è certi al 90% che gli effetti casuali non abbiano generato la differenza osservata. Dai dati della Tabella 1 e usando l'equazione (8), CL è ±1,64 x SE (Risparmi) per dimensioni di campione superiori a 120.

5.1 Valutazione delle interazioni di componenti multipli d'incertezza

Le equazioni (19) e (20) per la combinazione dei componenti d'incertezza possono essere utilizzate per stimare come gli errori di un componente influiranno sull'accuratezza del rendiconto totale dei risparmi. Potranno quindi essere messe in campo risorse di M&V per ridurre in modo redditizio l'errore nei risparmi rendicontati. Tali considerazioni progettuali dovranno tener conto dei costi e degli effetti sulla precisione dei risparmi di possibili miglioramenti nella precisione di ogni componente.
Le applicazioni software scritte per le funzioni dei comuni fogli di calcolo consentono una facile valutazione dell'errore netto associato alla combinazione di più componenti d'incertezza utilizzando le tecniche Monte Carlo. L'analisi Monte Carlo permette di valutare più scenari del tipo "Che cosa succederebbe se…" che rivelano un intervallo di possibili risultati, la loro probabilità di verificarsi e quale componente avrà il maggior effetto sul risultato finale. Tale analisi identifica i punti in cui devono essere assegnate risorse per il controllo degli errori. Qui di seguito viene presentata una semplice illustrazione dell'analisi "che cosa succederebbe se" relativa a un'AMEE per illuminazione. Un'apparecchiatura da 96 watt nominali viene sostituita con una da 64 watt nominali. Se detta apparecchiatura funziona per 10 ore al giorno, i risparmi annuali verrebbero calcolati come:

$$Risparmi\ annuali = \frac{(96-64)\times 10\times 365}{1.000} = 117kWh$$

La potenza del nuovo dispositivo di 64 watt è coerente e viene facilmente misurata con accuratezza. Tuttavia c'è molta differenza fra le potenze in watt della vecchia apparecchiatura e fra le ore di utilizzo in diverse ubicazioni. Le potenze in watt della vecchia apparecchiatura e le ore di utilizzo non sono facilmente misurate con certezza. Pertanto, anche i risparmi non saranno noti con certezza. La sfida nel progettare un piano di M&V è determinare l'impatto sui risparmi rendicontati se la misura dell'una o dell'altra di queste grandezze incerte è errata in quantità plausibili.

La Figura 2 mostra un'analisi di sensibilità dei risparmi per i due parametri, i watt della vecchia apparecchiatura e le ore di utilizzo. Ciascuno è variato fino al 30% ed è riportato l'impatto sui risparmi. Si può notare che i risparmi sono significativamente più sensibili alla variazione nella potenza in watt della vecchia apparecchiatura che non alle ore di utilizzo. Un errore di potenza in watt del 30% genera un errore di risparmi del 90%, mentre un errore del 30% in ore di funzionamento genera un errore di solo il 30% sui risparmi.

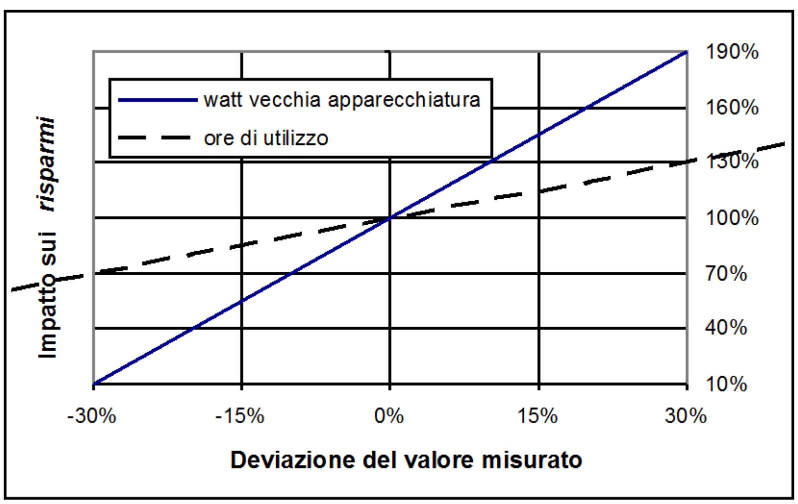

Figura 2 Esempio di analisi di sensibilità - Risparmi sull'illuminazione

Se il metodo di M&V proposto genera letture della potenza in watt della vecchia apparecchiatura con un intervallo d'incertezza di ± 5%, l'intervallo dei risparmi di energia elettrica sarà ± 15%. In altre parole, se la potenza in watt della vecchia apparecchiatura potesse essere compresa tra 91 e 101 watt, i risparmi potrebbero essere compresi tra 99 e 135 kWh all'anno. L'intervallo d'incertezza dei risparmi è di 36 kWh

(135 - 99). Se il valore marginale di energia elettrica è di 10 centesimi per kWh, l'intervallo d'incertezza sarà di circa $ 3,60 all'anno. Se la potenza in watt della vecchia apparecchiatura potesse essere stimata con maggiore precisione per molto meno di $ 3,60 all'anno, allora potrà valere la pena migliorare la misura con sforzi che dipendono dal numero d'anni di risparmi presi in considerazione.

La Figura 2 mostra che il termine delle ore di utilizzo ha un impatto inferiore sui risparmi finali in questo esempio (la linea delle ore di utilizzo è più piatta e indica sensibilità più bassa). È plausibile che l'errore di misura delle ore di funzionamento sia di ± 20%, per cui anche l'intervallo d'incertezza di risparmio di energia è ± 20% o ± 23 kWh (= 20% di 117 kWh). L'intervallo di risparmi è di circa 46 kWh (= 2 x 23 kWh), per un valore corrispettivo di $ 4,60 all'anno. Anche in questo caso può convenire aumentare la precisione nella misura delle ore di funzionamento se tale operazione può essere effettuata per molto meno di $ 4,60 all'anno, con sforzi che dipendono dal numero di anni di risparmio presi in esame.

L'intervallo dei possibili errori di risparmio da errori di misura delle ore d'esercizio (46 kWh) è maggiore di quella dell'errore di misura delle potenze in watt della vecchia apparecchiatura (36 kWh). Questo è l'effetto opposto a quello che ci si potrebbe attendere in base alla maggiore sensibilità dei risparmi rispetto alle ore di utilizzo, come si vede in Figura 2. La differenza è dovuta al fatto che l'errore plausibile nella misura delle ore d'esercizio (± 20%) è molto più grande dell'errore plausibile nella misura delle potenze in watt della vecchia apparecchiatura (± 5%).

L'analisi di sensibilità come sopra descritta può assumere molte forme. Il semplice esempio precedente intende illustrarne i principi. La simulazione Monte Carlo permette una valutazione complessa di molti parametri diversi, consentendo alla progettazione di un piano di M&V di concentrare le spese dove sono maggiormente necessarie per migliorare la precisione complessiva dei rendiconti di risparmio.

5.2 Determinazione di traguardi per l'incertezza quantificabile di campioni

Come discusso in precedenza, non tutte le incertezze possono essere quantificate. Tuttavia quelle che possono essere quantificate forniscono indicazioni per la progettazione dei piani di M&V. Considerando il costo di M&V dei vari approcci all'incertezza, il programma di M&V può generare il tipo di informazioni accettabile a tutti i lettori del rendiconto dei risparmi, compresi quelli che devono pagare i rendiconti di M&V. In ultima analisi qualsiasi piano di M&V dovrà rendicontare il livello atteso di incertezza quantificabile.

La determinazione dei risparmi di energia richiede la stima di una differenza nei livelli di energia piuttosto che la semplice misura del livello stesso d'energia. In generale nel calcolare una differenza per soddisfare un obiettivo di precisione relativa occorre una migliore precisione assoluta nelle misure dei componenti rispetto alla precisione assoluta richiesta dalla differenza. Ad esempio, supponiamo che il carico medio sia di circa 500 kW e che i risparmi previsti siano di circa 100 kW. Un errore di ± 10% con un criterio di confidenza del 90% ("90/10") può essere applicato in due modi:

- se applicato alle misure di carico, la precisione assoluta deve essere di 50 kW (10% di 500 kW) con percentuale di confidenza del 90%;
- se applicato ai risparmi rendicontati, la precisione assoluta dei risparmi deve essere di 10 kW (10% di 100 kW) allo stesso livello di confidenza del 90%. Per raggiungere questo obiettivo di precisione assoluta di 10 kW nei risparmi rendicontati occorrono precisioni assolute nella misura dei componenti di 7 kW (usando l'equazione (19) se entrambi i componenti devono avere la stessa precisione).

Chiaramente l'applicazione del criterio confidenza/precisione di 90/10 a livello dei risparmi richiede molta più precisione nella misura del carico rispetto ad un requisito di 90/10 a livello del carico.

Il criterio di precisione può essere applicato non solo ai risparmi di energia, ma anche ai parametri che determinano i risparmi. Ad esempio, si supponga che l'importo dei risparmi sia il prodotto del numero (N) di unità, ore (H) di funzionamento e variazione(C) in watt: Risparmi = N x H x C. Il criterio di 90/10 potrebbe venir applicato separatamente a ciascuno di questi parametri. Tuttavia il conseguimento di una precisione di 90/10 per ciascuno di questi parametri separatamente non implica che si ottenga una precisione di 90/10 per i risparmi, che è il parametro di interesse finale. Infatti utilizzando l'equazione (20), la precisione ad una confidenza del 90% sarebbe solo ± 17%. D'altra parte se il numero di unità e il cambiamento in watt si presumono noti senza errori, la precisione di 90/10 per le ore implica una precisione di 90/10 per i risparmi.

Lo standard di precisione potrebbe essere imposto a vari livelli. La scelta del livello di disaggregazione incide pesantemente sulla progettazione del piano di M&V e sui costi associati. In generale i requisiti di raccolta dei dati aumentano se i requisiti di precisione sono imposti a ciascun componente.
Se l'obiettivo principale è controllare la precisione dei risparmi per un progetto nel suo complesso, non è necessario imporre lo stesso requisito di precisione su ogni componente.

6 Esempio di analisi d'incertezza

Per illustrare l'uso dei vari strumenti statistici per l'analisi d'incertezza la Tabella 3 mostra un esempio di risultato dell'analisi di regressione di un foglio di calcolo. Si tratta di una regressione dei 12 valori mensili misurati dal contatore di energia elettrica di un edificio e dei gradi giorno di raffrescamento (GG_{raf}) nel corso di un periodo di un anno. Questo è solo un foglio di calcolo parziale. I valori specifici d'interesse sono evidenziati in corsivo.

Tabella 3 Esempio di risultati dell'analisi di regressione di un foglio di calcolo

RIEPILOGO DEI RISULTATO

Statistica di regressione					
R multiplo	0,97				
R quadrato	0,93				
R quadrato adeguato	0,92				
Errore standard	367,50				
Osservazioni	12,00				
	Coefficienti	*Errore standard*	*T-stat*	*95% inf.*	*95% sup.*
Intercetta	5.634,15	151,96	37,08	5.295,56	5.972,74
GG_{raf}	7,94	0,68	11,64	6,42	9,45

Per un riferimento di 12 valori mensili di kWh e GG_{raf} il modello di regressione derivato è:

Consumo mensile di elettricità = 5.634,15 + (7,94 x GG_{raf})

Il coefficiente di determinazione, R^2, (indicato come "R quadrato" in Tabella 3) è molto alto pari a 0,93, che indica che il 93% della variazione dei 12 dati di consumo di energia viene spiegato dal modello utilizzando i dati GG_{raf}. Questo fatto implica una relazione molto stretta e che il modello può essere utilizzato per stimare i termini di aggiustamento del pertinente modello di risparmio = (Consumo o domanda del periodo di riferimento – Consumo o domanda del periodo di rendicontazione) ± Aggiustamenti. Il coefficiente stimato di 7,94 kWh per GG_{raf} ha un errore standard di 0,68. Questo SE porta a un t di Student (indicato come "T stat" in Tabella 3) di 11,64. Tale t di Student viene quindi confrontato con il valore critico appropriato della Tabella 1 (t = 2,2 per 12 punti dati e 95% di confidenza). Dato che 11,64 supera 2,2, GG_{raf} è una variabile indipendente altamente significativa. Il foglio di calcolo mostra anche che l'intervallo per il coefficiente a livello di confidenza del 95% va da 6,42 a 9,45 e implica una precisione relativa di ± 19% (= (7,94 - 6,42) / 7,94). In altre parole, abbiamo una percentuale di confidenza del 95% che ogni ulteriore GG_{raf} aumenti il consumo di kWh da 6,42 a 9,45 kWh.

L'errore standard della stima utilizzando la formula di regressione è 367,5. I GG_{raf} medi mensili sono di 162 (valore non mostrato in uscita). Questo valore GG_{raf} viene inserito nel modello di regressione come esempio per prevedere quale sarebbe stato il consumo elettrico in condizioni medie di raffrescamento:

Consumo presunto = 5.634 + (7,94 x 162) = 6.920 kWh per mese con valore medio mensile di GG_{raf}

Utilizzando un valore *t* della Tabella 1 di 2,2, per 12 punti dati e per un livello di confidenza del 95%, l'intervallo delle possibili previsioni è:

Intervallo di previsioni = 6.920 ± (2,2 x 367,5) = da 6.112 a 7.729 kWh.

La precisione assoluta è di circa ± 809 kWh (= 2,2 x 367,5) e la precisione relativa è invece di ± 12% (= 809/6.920). Il valore descritto del foglio di calcolo per l'errore standard della stima ha fornito le informazioni necessarie per calcolare la precisione relativa attesa dall'uso del modello di regressione per qualsiasi mese, in questo caso 12%. Se il consumo del periodo di rendicontazione è stato di 4.300 kWh, i risparmi sono calcolati come: Risparmi = (6.920 – 4.300) = 2.620 kWh. Poiché per ottenere il valore di energia elettrica del periodo di rendicontazione è stato usato il contatore del distributore, il suo valore rendicontato può essere trattato come esatto al 100% (SE = 0%) dato che il contatore del distributore definisce gli importi pagati, indipendentemente dall'errore strumentale. Lo SE del risparmio di energia sarà:

$$SE(\textit{risparmi mensili}) = \sqrt{SE(\textit{consumo riferimento adeguato})^2 + SE(\textit{consumo rendicontazione})^2}$$

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

Usando un t di 2.2, l'intervallo dei possibili risparmi mensili è

Intervallo di risparmi = 2.620 ± (2,2 x 367,5) = 2.620 ± -810 = da 1.810 a 3.430

Per determinare la precisione del totale annuo dei risparmi mensili, si presume che l'errore standard dei risparmi di ogni mese sarà lo stesso. Allora i risparmi annui riportati hanno un errore standard di:

$$SE\ (\textit{risparmi annuali}) = \sqrt{12 \times 367{,}5^2} = 1.273\ \text{kWh}$$

Poiché t deriva dal modello di riferimento, rimane pari al valore 2,2 utilizzato sopra. Pertanto la precisione assoluta nei risparmi annui è di 2,2 x 1.273 = 2.801 kWh. Supponendo uguali risparmi mensili di 2.620 kWh, i risparmi annuali sono di 31.440 kWh e la precisione relativa del rendiconto annuale dei risparmi è del 9% (= (2.801 / 31.440) x 100).

EVO ringrazia le organizzazioni sottoscrittrici:

Organizzazioni Principali:

BC Hydro
Services Industriels de Genève

Organizzazioni Senior:

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

Grandi organizzazioni Governative:

Ontario Power Authority
Bonneville Power Administration

Piccole organizzazioni Governative:

ADENE – Agência para a Energia(Portuguese Energy Agency)

Settore Educativo:

Université de Genève

Organizzazioni associate dell'Allegato 1:

Navigant Consulting Inc.

Quantum Energy Services & Technologies, Inc. (QuEST)

EEVS – Energy Efficiency Verification Specialists

HEP-ESCO d.o.o.

Organizzazioni Non Profit Non Allegato 1:

Taiwan Green Productivity Foundation (TGPF)