



Inferenza bayesiana per l'analisi dei dati di prove di fatica

A. De Iorio, F. Penta, P. Pinto

Università degli Studi di Napoli Federico II, Dipartimento di Progettazione e Gestione Industriale, p.le V. Tecchio n° 80, 80125, NAPOLI, antdeior@unina.it.

M. Guida

Università degli Studi di Salerno, Dipartimento di Ingegneria dell'Informazione ed Ingegneria Elettrica, via Ponte don Melillo n°1, 84084, Fisciano (SA), mguida@unisa.it.

RIASSUNTO. Si prende in esame il problema della caratterizzazione a fatica dei materiali mediante campioni di dati poco numerosi. Se ne prospetta una soluzione che impiega la metodologia bayesiana. Per la vasta famiglia degli acciai al carbonio, laminati a caldo o bonificati che siano, utilizzando dati di letteratura ed alcune consolidate correlazioni tra proprietà di fatica e resistenza statica, è definita una funzione di densità di probabilità *a priori* in grado di condensare gran parte delle informazioni disponibili. Queste ultime, in uno con quelle fornite dalla sperimentazione diretta, da esaminare mediante il teorema di Bayes, permettono di identificare con grande accuratezza la resistenza a fatica del particolare acciaio provato. L'efficacia del metodo proposto è verificata con una sperimentazione virtuale su un ipotetico acciaio condotta con il metodo Montecarlo.

ABSTRACT. The problem of fatigue characterization of materials by small data samples is here considered and a Bayesian procedure for data analysis is presented. For both hot rolled and quenched and tempered carbon steels, on the basis of some literature data and by well-established correlations between fatigue and tensile properties, a prior probability density function, able to model most of the available information on material fatigue strength, is defined. Prior information, when fused by means of Bayes theorem with information from experimental data, allows to identify with greater accuracy the fatigue properties of the tested steel. A Montecarlo study is carried out to verify the effectiveness of the proposed method, generating small samples of fatigue data by virtual testing.

PAROLE CHIAVE: curve PSN, inferenza bayesiana, *fatigue design curve*.

NOMENCLATURA

A	coefficiente di durata;	N_i	durata dell'i-esima prova;
A_1, A_2	stime dell'intercetta e della pendenza, rispettivamente, della retta di regressione $\delta - b$;	$\hat{\sigma}^2$	stima puntuale frequentista di σ^2 ;
B	supporto della funzione di densità <i>a posteriori</i> π_1 ;	S_i	ampiezza di carico dell'i-esima prova;
B_1, B_2	stime dell'intercetta e della pendenza della retta di regressione $\log UTS - \log S'_f$, rispettivamente;	S'_f, b	coefficiente di resistenza a fatica ed esponente di fatica, rispettivamente;
D	costante di normalizzazione;	S_N	resistenza a fatica ad N cicli;
e_i	variabile normale standard;	$t_{(l)}$	variabile di Student con l gradi di libertà;
E	modulo di Young;	$t_{(\eta, 1-\gamma, l)}$	$(1-\gamma)$ -esimo percentile della distribuzione di Student non centrale, con l numero di gradi di libertà ed η parametro di non centralità;
$f(A \alpha, \beta, \sigma)$	funzione di verosimiglianza (<i>likelihood</i>) del campione di dati;	UTS	ultimate tensile stress;
$f_1(b \delta), f_2(\delta UTS)$	funzioni di densità di probabilità a priori di $b \delta$ e δUTS rispettivamente;	ε_i	errore random;
$\log UTS, \bar{\delta}$	medie dei campioni di $\log UTS$ e di δ , rispettivamente;	x_i	variabile indipendente del modello di regressione lineare;
$l_{ti}(y_x^p)$	limite inferiore di tollerabilità frequentista;	y_i	durata trasformata dell'i-esima prova;
m	numerosità del campione di acciai;	y_x^p	p -quantile delle durate trasformate;
n	numerosità del campione di dati di fatica;	$y_{x, 1-\gamma}^p$	limite inferiore di tollerabilità bayesiano;
		z_p	p -quantile della distribuzione normale standard;
		α, β	intercetta e pendenza, rispettivamente, della

	retta di regressione;		trasformata media;
$\bar{\alpha}, \bar{\beta}$	variabili dummy;	$\pi(\alpha, \beta UTS)$	funzione di densità di probabilità <i>a priori</i>
$\hat{\alpha}, \hat{\beta}$	stime puntuali frequentiste rispettivamente di α e β ;		congiunta di α e β , dato <i>UTS</i> ;
$\hat{\alpha}_B, \hat{\beta}_B$	stime puntuali bayesiane, rispettivamente di α e β ;	$\pi_1(\alpha, \beta, \sigma \Delta, UTS)$	funzione di densità di probabilità <i>a posteriori</i> congiunta di α, β e σ ;
$\delta = \log(S'_f / UTS)$	coefficiente adimensionale di resistenza a fatica trasformato;	$\pi_2(\alpha, \beta \Delta, UTS)$	funzione di densità di probabilità <i>a posteriori</i> marginale congiunta di α e β ;
Δ	campione di dati;	$\pi_3(y_x^p \Delta, UTS, x)$	funzione di densità di probabilità <i>a posteriori</i> di y_x^p ;
$\Delta\sigma, \Delta\varepsilon, \Delta\varepsilon_{el}$	<i>stress range, strain range</i> ed <i>elastic strain range</i> , rispettivamente;	$\pi_4(\bar{\alpha}, \bar{\beta}, y_x^p \Delta, UTS, x)$	funzione di densità di probabilità <i>a posteriori</i> congiunta di α, β ed y_x^p ;
$\hat{\mu}_f, \hat{s}_f^2$	stime puntuali della media e della varianza di $\log S'_f \log UTS$;	σ^2	varianza delle durate trasformate;
$\hat{\mu}_{\nu B}$	stima puntuale bayesiana della durata	$\sigma_{b \delta}^2, \sigma_f^2$	varianze delle funzioni di densità di probabilità <i>a priori</i> di $b \delta$ e $\log S'_f \log UTS$ rispettivamente;

1. INTRODUZIONE

Sebbene sia pratica comune studiare il fenomeno della fatica eseguendo prove in condizioni strettamente controllate, i dati che si ottengono sono, com'è noto, sempre notevolmente dispersi. Il che giustifica il frequente ricorso ad una loro analisi statistica, non solo per definire le proprietà di resistenza a fatica del materiale tramite un modello probabilistico, ma anche per costruire una *fatigue design curve*, che tenga conto, in modo conservativo, sia dell'incertezza prodotta dalla natura aleatoria del fenomeno, sia della non perfetta conoscenza dei parametri del modello dovuta alla dimensione comunque finita del campione di dati [1-4]. La metodologia di analisi più impiegata si avvale degli strumenti e dei metodi della teoria frequentista, che però conduce a risultati di scarsa utilità tecnica quando il campione di dati è poco numeroso. Questa circostanza si verifica, tipicamente, nelle prime fasi di sviluppo di un nuovo prodotto e/o componente, cioè quando sono incognite le proprietà tecniche e commerciali del prodotto e non sono giustificabili gli oneri di una estesa e *time-consuming* attività sperimentale di caratterizzazione.

Un modo molto efficace di procedere, per superare tale inconveniente, sembra il ricorso alla metodologia di analisi bayesiana [5], che offre la possibilità di utilizzare, insieme ai dati sperimentali, anche informazioni prodotte con precedenti studi teorici e/o analoghe indagini sperimentali. Queste ultime, in linea di principio, devono essere opportunamente formalizzate in termini di funzioni di densità di probabilità *a priori*, per i parametri del modello probabilistico, e combinate, mediante il teorema di Bayes, con le informazioni contenute nel campione di dati.

Una conferma dei vantaggi ottenibili impiegando la

metodologia bayesiana sono i risultati conseguiti nello studio delle proprietà di *crack propagation* e di tollerabilità del danno di alcuni componenti meccanici [6-8]. Il suo utilizzo per caratterizzare le proprietà di resistenza a fatica del materiale, tuttavia, risulta molto raro, a giudicare dai pochissimi lavori dedicati all'argomento, quali, ad es., il lavoro di Madsen [9], che ha analizzato il comportamento di alcuni giunti di strutture off-shore utilizzando informazioni *a priori* dedotte dal comportamento di giunti simili, e quello di Edwards et Al. [10], che, adottando distribuzioni *a priori* non informative, hanno superato il problema dell'analisi dei campioni censurati di piccole dimensioni. Probabilmente, la scarsa attenzione rivolta finora alla metodologia bayesiana è dovuta, da una parte, al senso di diffidenza con cui è accolto il concetto di probabilità soggettiva, che è alla base del metodo, e, dall'altra, ai maggiori oneri di calcolo solitamente necessari rispetto ai metodi frequentisti. Per questo si è ritenuto opportuno indagare sul tema della definizione di una procedura bayesiana per l'analisi di dati di fatica ponendo l'accento proprio su questi due aspetti del problema.

Ricorrendo al modello lineare normale con varianza uniforme, generalmente adottato per definire il legame carichi-durate delle leghe strutturali, viene definito un modello bayesiano per le proprietà di resistenza a fatica degli acciai al carbonio. Le funzioni di densità *a priori* sono costruite analizzando alcuni dati di letteratura e formalizzando in ambito bayesiano alcune correlazioni ben note tra proprietà di fatica e resistenza statica. Esso conduce a stime puntuali dei parametri di interesse ed alla definizione di una *design curve* utilizzando comuni strumenti di calcolo. Con delle simulazioni numeriche vengono infine evidenziati i vantaggi offerti, in termini di precisione delle stime puntuali e per intervalli, dal metodo di analisi bayesiano proposto rispetto a quello frequentista.

2. MODELLO

Le ipotesi di partenza dell'analisi sono quelle classiche, generalmente valide per gli acciai e per altre leghe strutturali, compendiate nel modello probabilistico:

$$N_i = A \cdot S_i^{-\beta} \cdot \varepsilon_i \quad (1)$$

in cui N_i (con $i=1 \dots n$) è la durata osservata dell' i -esima prova condotta con un carico alterno-simmetrico di ampiezza S_i ; A e β sono parametri del materiale, mentre ε_i è un errore random di tipo moltiplicativo che è distribuito come una variabile log-normale con varianza indipendente da S_i .

Ai fini dell'analisi, è conveniente analizzare il legame stabilito dall'Eq.1, in termini di durate trasformate $y_i = \log(N_i)$ e di ampiezze trasformate $u_i = \log(S_i)$, ed eseguire il cambiamento di variabile $u_i = x_i + \bar{u}$, ove $\bar{u} = (1/n) \cdot \sum_{i=1}^n x_i$ è la media delle ampiezze trasformate con cui sono state eseguite le n prove, così da ottenere il semplice modello di regressione lineare:

$$y_i = \alpha - \beta x_i + \sigma e_i \quad (2)$$

in cui $\alpha = \log(A) - \beta \cdot \bar{u}$ ed $e_i \sim N(0,1)$.

Nelle ipotesi poste, dunque, le durate trasformate sono variabili random normali ed indipendenti, con media $\mu_{x_i} = \alpha - \beta \cdot x_i$ e varianza σ^2 .

Per stimare i parametri dell'Eq.2 è necessario ricorrere ad un'analisi statistica del campione di dati, Δ , ottenuto dalla sperimentazione.

Nel presente approccio, si prende in esame il caso di campioni non censurati, costituiti da n coppie del tipo (x_i, y_i) . Le informazioni del campione riguardanti i parametri del modello sono introdotte nell'analisi bayesiana attraverso la seguente funzione di verosimiglianza:

$$f(\Delta | \alpha, \beta, \sigma) \propto \frac{1}{\sigma^n} \cdot \exp \left\{ -\frac{1}{2} \left[\sum_{i=1}^n (y_i - \alpha + \beta \cdot x_i)^2 \right] \right\} \quad (3)$$

che esprime la probabilità di osservare il set di dati sperimentali, noti che siano i parametri α , β e σ . Attraverso semplici manipolazioni algebriche riportate in appendice, la Eq.3 può assumere la forma più compatta riportata nell'Eq.3a, che meglio si presta all'implementazione numerica:

$$f(\Delta | \alpha, \beta, \sigma) \propto \frac{1}{\sigma^n} \cdot \exp \left\{ -\frac{1}{2} \cdot \frac{(n-2) \cdot \hat{s}^2}{\sigma^2} \right\} \cdot \exp \left\{ -\frac{1}{2} \frac{n \cdot (\alpha - \hat{\alpha})^2 + \left(\sum_{i=1}^n x_i^2 \right) \cdot (\beta - \hat{\beta})^2}{\sigma^2} \right\} \quad (3a)$$

in cui

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^n y_i, \quad \hat{\beta} = -\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \quad \text{ed} \quad \hat{s}^2 = \frac{\sum_{i=1}^n (y_i - \hat{\alpha} + \hat{\beta} \cdot x_i)^2}{n-2}$$

sono gli stimatori frequentisti, non distorti e di varianza minima, dei parametri α , β e della varianza σ^2 , rispettivamente.

3. INFORMAZIONI A PRIORI

Le informazioni disponibili all'analista prima che osservi i dati, da formalizzare ed introdurre nell'inferenza mediante la distribuzione a priori, possono essere di varia natura, in quanto possono: - scaturire da considerazioni di carattere fisico o tecnologico; - dedursi da modelli micro-meccanici più raffinati; - consistere nel giudizio personale di uno o più esperti del settore; - essere ricavate dall'analisi di dati storici relativi a problemi analoghi a quello in esame.

Con riferimento al modello (2), in particolare, le informazioni minime disponibili, derivanti da dati storici, possono così sintetizzarsi:

- $\beta \in]0, +\infty[$, essendo sempre $\frac{d \log(S_N)}{d \log(N)} < 0$,
- $\alpha > \log(1/2) + \beta \cdot [\log(UTS) - \bar{u}]$

Questa seconda relazione, deriva dalla considerazione che nel campo della LCF la curva di fatica degli acciai ha tipicamente la concavità rivolta verso il basso e tende a raggiungere la posizione limite, cui corrisponde

$$\frac{S_N}{UTS} \rightarrow 1 \quad \text{per} \quad N \rightarrow \frac{1}{2}$$

Nell'analisi bayesiana delle proprietà di HCF degli acciai, questo stato di conoscenze riguardanti α e β dedotto da dati storici, se formalizzato mediante distribuzioni *a priori*, porterà a stime *a posteriori* sicuramente più realistiche di quelle ottenibili con le seguenti distribuzioni non informative:

$$\pi(\alpha) \propto \text{cost}, \quad \pi(\beta) \propto \text{cost}, \quad \text{con} \quad \alpha, \beta \in \mathbb{R} \quad (4)$$

cui si ricorre quando non si dispone di alcuna informazione per α e β . Si osservi, inoltre, che, se unitamente alle densità a priori (4), anche lo stato di conoscenza *a priori* riguardo la deviazione standard σ è formalizzato mediante l'analoga distribuzione non informativa, $\pi(\sigma) \propto \sigma^{-1}$, l'inferenza bayesiana conduce a stime dei parametri α , β e σ numericamente coincidenti con quelle frequentiste.

Costruzione della funzione di densità a priori

Per esemplificare la costruzione di una funzione di densità *a priori*, si è considerato il caso degli acciai al carbonio, per il quale è possibile definire una distribuzione *a priori* congiunta per α e β di tipo informativo, deducendola dai valori di letteratura del coefficiente di resistenza a fatica, S'_f , e dell'esponente di fatica, b , della formula di Basquin:

$$\frac{\Delta\sigma}{2} = \frac{E \cdot \Delta\varepsilon_{el}}{2} = S'_f \cdot (2N_f)^{-b}$$

che esprime lo *stress-range*, $\Delta\sigma$, o l'equivalente *range* di deformazione elastica, $\Delta\varepsilon_{el} = \Delta\sigma/E$, con E modulo di Young, in grado di portare a rottura il provino in N_f cicli.

E' da premettere che nel campo dell'*HCF* ($N_f \geq 10^5$) il range di deformazione plastica è notoriamente molto più piccolo di quello totale e di quello elastico ed è ipotizzabile che le modalità di controllo, di carico o di deformazione, delle prove non influiscano significativamente sulle durate, per cui, in prima approssimazione, si può scrivere:

$$S_N \approx E \cdot \frac{(\Delta\varepsilon)_N}{2} \approx E \cdot \frac{(\Delta\varepsilon_e)_N}{2} = S'_f \cdot (2N_f)^{-b}$$

il che autorizza a trasferire le informazioni raccolte per S'_f e b ai parametri α e β del modello in esame, relativo invece a prove condotte in controllo di carico.

I risultati dei numerosi studi finora svolti per stabilire delle relazioni empiriche tra le proprietà tensili e quelle di fatica degli acciai [11÷14, ad es.] concordano circa

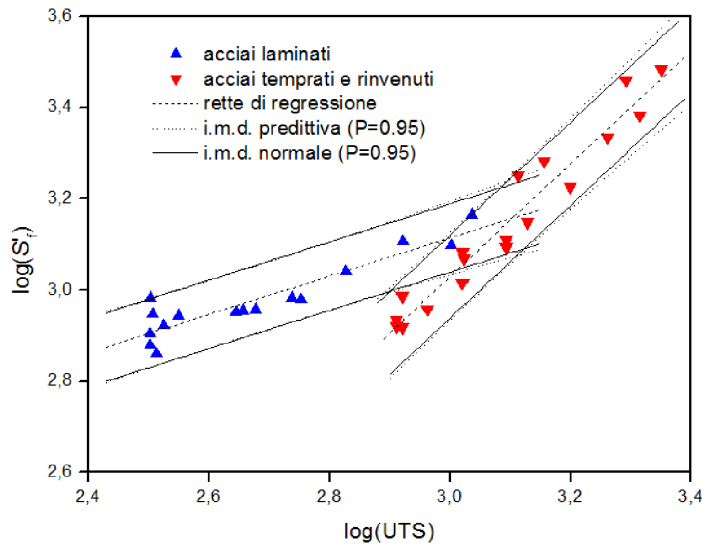


Figura 1: Legame tra il coefficiente di resistenza a fatica, S'_f , e il carico di rottura, UTS , per gli acciai al carbonio ed intervalli di massima densità (i.m.d.) *a priori*, con $P=0.95$.

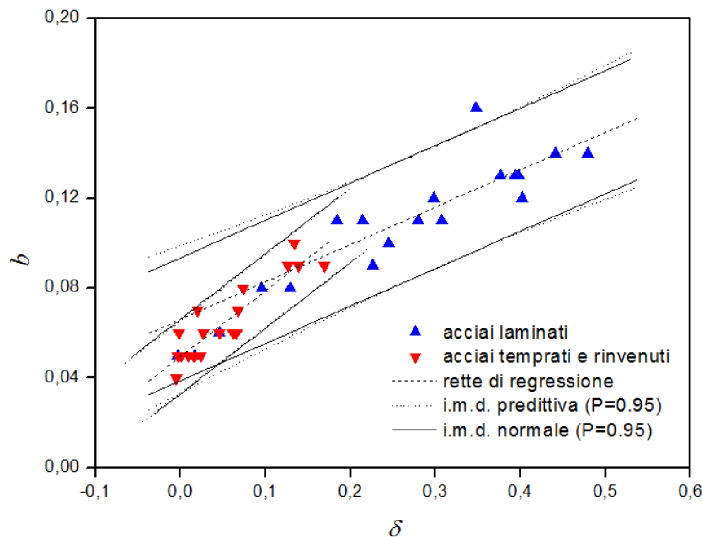


Figura 2: Legame b - δ per gli acciai al carbonio ed intervalli di massima densità (i.m.d.) *a priori*, con $P=0.95$.

l'esistenza di una dipendenza lineare tra il $\log(UTS)$ ed il $\log(S'_f)$. I dati [15] riportati nei diagrammi delle Fig. 1 e 2, oltre a confermare l'esistenza di questo legame, evidenziano anche una correlazione tra b ed il parametro $\delta = \log(S'_f / UTS)$

Stimati, allora, con un'analisi bayesiana non informativa, i coefficienti \hat{A}_1 e \hat{A}_2 della retta di regressione normale di Fig. 2, nonché la varianza di $b, s_{b|\delta}^2$, intorno a tale retta, lo stato di conoscenza riguardo a b , per un acciaio di cui sia noto il valore di δ , è esprimibile mediante la seguente distribuzione predittiva di $b | \delta$, secondo cui:

$$\frac{b - \hat{A}_1 - \hat{A}_2 \cdot \delta}{s_{b|\delta} \cdot \sqrt{c_{b|\delta}}} \sim t_{(m-2)}$$

dove $t_{(m-2)}$ è una variabile di Student con $m-2$ gradi di libertà, essendo m il numero di dati (δ_i, b_i) presenti nel campione, e

$$c_{b|\delta} = 1 + \frac{1}{m} + \frac{(\delta - \bar{\delta})^2}{\sum_{i=1}^m (\delta_i - \bar{\delta})^2}$$

con

$$\bar{\delta} = \frac{1}{m} \cdot \sum_{i=1}^m \delta_i$$

La scelta della funzione di densità predittiva è infatti vincolata alla necessità di tener conto, con l'informazione a priori riguardo a $b | \delta$, anche della variabilità di b tra gli individui della popolazione di acciai in esame, oltre che della incertezza delle stime \hat{A}_1 , \hat{A}_2 e $s_{b|\delta}^2$. E' conveniente, però, approssimare la

densità predittiva con la distribuzione normale $f_2(b | \delta)$, avente media $\hat{A}_1 + \hat{A}_2 \cdot \delta$ e varianza $\sigma_{b|\delta}^2$ indipendente da δ , ponendo:

$$\sigma_{b|\delta}^2 = s_{b|\delta}^2 \cdot c_{b,max}$$

ove $c_{b,max}$ è il valore massimo attinto da $c_{b|\delta}$ nel range di variabilità dei dati.

Similmente, in virtù del legame esistente tra $\log S'_f$ e

$\log UTS$ e sulla base delle stesse considerazioni già svolte per $b | \delta$, lo stato di conoscenze *a priori* riguardo $\log S'_f$, sempre per un acciaio al carbonio, nota che sia la sua UTS , può essere approssimato mediante la funzione di densità normale avente media:

$$\hat{\mu}_f = \hat{B}_1 + \hat{B}_2 \cdot \log(UTS)$$

con \hat{B}_1 e \hat{B}_2 stime dei coefficienti della retta di regressione normale di Fig. 1, e varianza:

$$\sigma_f^2 = \hat{s}_f^2 c'_{max}$$

ove \hat{s}_f^2 è la stima puntuale della varianza di $\log S'_f | \log UTS$ e $c'_{max} = \max_{\log UTS} (c')$, essendo

$$c' = 1 + \frac{1}{m} + \frac{(\log UTS - \overline{\log UTS})^2}{\sum_{i=1}^m (\log UTS_i - \overline{\log UTS})^2}$$

con

$$\overline{\log UTS} = \frac{1}{m} \cdot \sum_{i=1}^m \log UTS_i$$

Ne segue, quindi, che, se di un dato acciaio al carbonio è nota l' UTS , la distribuzione *a priori* $f_2(\delta | UTS)$, per il parametro $\delta = \log(S'_f / UTS)$ è anch'essa normale, con media

$$\mu_{\delta UTS} = \hat{\mu}_f - \log UTS = \hat{B}_1 - (1 - \hat{B}_2) \cdot \log UTS$$

e varianza $\sigma_{\delta UTS}^2 = \sigma_f^2$, per cui la distribuzione *a priori* congiunta di b e δ , dato UTS , sarà data dalla distribuzione normale bivariata (Eq. 5).

I parametri α e β del modello (2) sono però legati ai valori di b e δ dalle seguenti relazioni:

$$\begin{cases} \alpha = \frac{1}{b} \cdot [\delta - \bar{u} + \log(S_u)] - \log(2) \\ \beta = \frac{1}{b} \end{cases} \quad (6)$$

che permettono di ricavare, mediante il teorema di trasformazione, la loro distribuzione a priori da quella di b e δ (Eq. 5a), essendo β^{-3} il determinante jacobiano della trasformazione definita dalle Eq.6. I valori

$$f(b, \delta | UTS) = f_1(b | \delta) \cdot f_2(\delta | UTS) =$$

$$= \frac{1}{2\pi \cdot \sigma_{b|\delta} \cdot \sigma_{\delta UTS}} \exp \left\{ -\frac{1}{2} \left[\frac{b - \hat{A}_1 - \hat{A}_2 \cdot \delta}{\sigma_{b|\delta}} \right]^2 - \frac{1}{2} \left[\frac{\delta - \hat{B}_1 + (1 - \hat{B}_2) \cdot \log UTS}{\sigma_{\delta UTS}} \right]^2 \right\} \quad (5)$$

$$\begin{aligned} \pi(\alpha, \beta | UTS) &= \frac{\beta^{-3}}{2\pi \sigma_{b|\delta} \sigma_{\delta UTS}} \cdot \exp \left\{ -\frac{1}{2} \left[\frac{1/\beta - \hat{A}_1 - \hat{A}_2 [(\alpha + \log 2) / \beta + \bar{u} - \log UTS]}{\sigma_{b|\delta}} \right]^2 \right\} \times \\ &\times \exp \left\{ -\frac{1}{2} \left[\frac{(\alpha - \log 2) / \beta + \bar{u} - \hat{B}_1 + \hat{B}_2 \cdot \log UTS}{\sigma_{\delta UTS}} \right]^2 \right\} \end{aligned} \quad (5a)$$

numerici dei parametri della funzione $\pi(\alpha, \beta | UTS)$, calcolati sulla base dei dati sperimentali raccolti, sono riportati nella Tab. 1.
 Nel diagramma di Fig. 3 sono invece mostrate le curve di livello delle funzioni di densità *a priori* $\pi(\alpha, \beta | UTS)$ (adimensionalizzate entrambe rispetto alle loro densità modali), relative a due acciai laminati aventi *UTS* rispettivamente pari a 350 e 1000 MPa, da caratterizzare con prove tali da aversi $\bar{u} - \log(UTS) = 0.427$. Dal loro esame si evince come la conoscenza del tipo di acciaio e della sua *UTS* consenta di definire, a partire dalle informazioni dei dati di Fig. 1 e Fig. 2, funzioni di densità *ad hoc* per il materiale da caratterizzare. Per la deviazione standard σ , infine, non è stato possibile

rintracciare in letteratura informazioni sufficienti per definire una distribuzione *a priori* uni-modale. Per tale motivo, lo stato di conoscenza *a priori* per σ è prudenzialmente espresso mediante la funzione di densità non informativa $\pi(\sigma) \propto \sigma^{-1}$.

4. INFERENZA A POSTERIORI

Le informazioni *a priori* sui parametri del modello possono essere combinate con quelle del campione di dati sperimentali mediante il teorema di Bayes, per ottenere la distribuzione *a posteriori* dei parametri riportata in Eq. 7.

$$\pi_1(\alpha, \beta, \sigma | \Delta, UTS) = \frac{f(\Delta | \alpha, \beta, \sigma) \cdot \pi(\alpha, \beta | UTS) \cdot \sigma^{-1}}{\iiint_B f(\Delta | \alpha, \beta, \sigma) \cdot \pi(\alpha, \beta | UTS) \cdot \sigma^{-1} d\alpha d\beta d\sigma} = \tag{7}$$

$$= \frac{\sigma^{-(n+1)}}{D} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \alpha + \beta \cdot x_i)^2\right\} \cdot \pi(\alpha, \beta | UTS)$$

in cui

$$D = \iiint_B \sigma^{-n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \alpha + \beta x_i)^2\right\} \frac{\pi(\alpha, \beta | UTS)}{\sigma} d\alpha d\beta d\sigma$$

con $B = \mathbb{R}^2 \times]0, +\infty]$, è il fattore di normalizzazione.

	\hat{A}_1	\hat{A}_2	$\hat{\sigma}_{\beta}$	\hat{B}_1	\hat{B}_2	$\hat{\sigma}_{\beta UTS}$
ACCIAI LAMINATI	6,610 10 ⁻²	1,665 10 ⁻¹	1,453 10 ⁻²	1,858	4,189 10 ⁻¹	3,826 10 ⁻²
ACCIAI TEMPRATI E RINVENUTI	4,950 10 ⁻²	2,924 10 ⁻¹	8,491 10 ⁻³	-6,794 10 ⁻¹	1,236	4,901 10 ⁻²

Tabella 1: Parametri della funzione di densità *a priori* per gli acciai al carbonio.

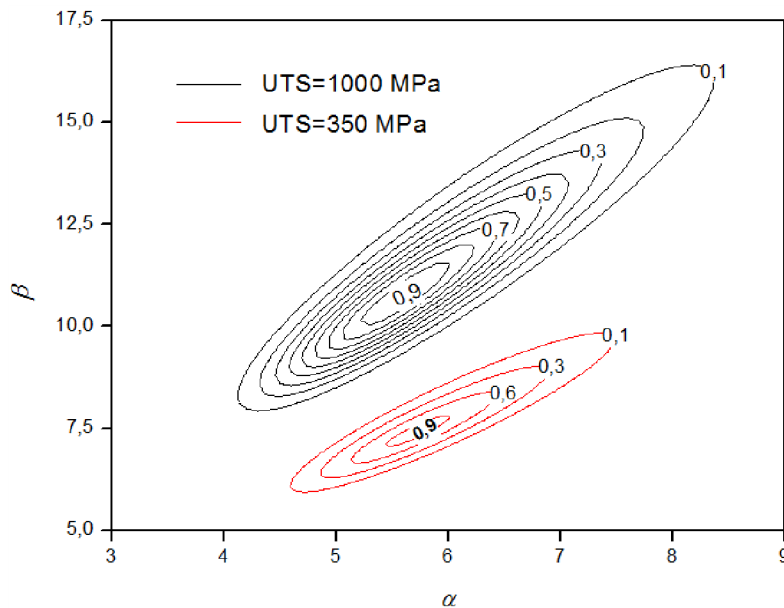


Figura 3: Funzioni di densità *a priori* (adimensionalizzate) per i parametri α e β di due acciai al carbonio laminati.

Si noti che, grazie alla particolare scelta della funzione di densità *a priori* di σ , il calcolo del fattore di normalizzazione è parzialmente semplificato, poiché è possibile eseguire l'integrazione rispetto a σ in forma chiusa, ottenendo l'Eq. 8.

Per lo stesso motivo, è possibile dedurre dalla Eq.4 l'espressione della funzione di densità marginale *a posteriori* congiunta dei parametri α e β , riportata nella Eq. 9, con la quale valutare le stime puntuali bayesiane $\hat{\alpha}_B$, $\hat{\beta}_B$ e $\hat{\mu}_{xB}$ rispettivamente di α , β e delle durate trasformate medie μ_x :

$$\hat{\alpha}_B = E\{\alpha\} = \iint \alpha \cdot \pi(\alpha, \beta | \Delta, UTS) \cdot d\alpha d\beta$$

$$\hat{\beta}_B = E\{\beta\} = \iint \beta \cdot \pi(\alpha, \beta | \Delta, UTS) \cdot d\alpha d\beta$$

$$\hat{\mu}_{xB} = E\{\mu_x\} = \hat{\alpha}_B + \hat{\beta}_B \cdot x$$

Queste sono, coerentemente con lo stato di conoscenza *a posteriori*, le migliori stime possibili nel senso dell'errore quadratico, dal momento che uno stimatore del tipo $\hat{z} = E\{z\}$, con z generica variabile aleatoria, ha la proprietà di ridurre al minimo (pari alla varianza di z) il valore atteso di $(z - \hat{z})^2$.

Per la distribuzione *a posteriori*, $\pi_3(y_x^p | \Delta, UTS, x)$, del generico p -quantile delle durate trasformate ottenibili al livello di carico x :

$$y_x^p = \alpha - \beta \cdot x + \sigma \cdot z_p$$

con z_p p -quantile della distribuzione normale standard, conviene effettuare il seguente cambiamento di variabili:

$$\begin{cases} \bar{\alpha} = \alpha, \\ \bar{\beta} = \beta, \\ y_x^p = \alpha - \beta x + \sigma z_p \end{cases} \quad (6)$$

ed applicare il teorema di trasformazione, ottenendo così dalla Eq.4 la densità congiunta *a posteriori* di $(\bar{\alpha}, \bar{\beta}, y_x^p)$:

$$\pi_4(\bar{\alpha}, \bar{\beta}, y_x^p | \Delta, UTS, x) = |J^{-1}| \cdot \pi_1\left(\bar{\alpha}, \bar{\beta}, \frac{y_x^p - \alpha - \beta x}{z_p} | \Delta, UTS\right)$$

in cui $J = z_p$ è il determinante jacobiano della trasformazione definita dalle Eq.6.

La funzione $\pi_3(y_x^p | \Delta, UTS, x)$ si ottiene marginalizzando

rispetto alle variabili $\bar{\alpha}$ e $\bar{\beta}$.

Le informazioni contenute nella distribuzione $\pi_3(y_x^p | \Delta, UTS, x)$ consentono non solo di effettuare la stima puntuale *a posteriori* delle durate aventi probabilità $1-p$ di essere superate al generico livello di carico x , ma anche di tracciare per punti la *design curve* bayesiana, risolvendo per assegnati valori di x la seguente equazione nell'incognito limite di tollerabilità bayesiano $y_{x,1-\gamma}^p$:

$$\int_{-\infty}^{y_{x,1-\gamma}^p} \pi(y_x^p | \Delta, UTS, x) \cdot dy_x^p = 1 - \gamma \quad (10)$$

ove $1 - \gamma$ è il livello di credibilità.

La *design curve* frequentista è invece definita mediante un'equazione analoga alla Eq.10, nella quale in luogo della densità *a posteriori* di y_x^p comparirà la distribuzione campionaria di un opportuno stimatore del p -quantile. Vale la pena sottolineare, però, che l'analogia tra le due curve è solo formale, in quanto, al di là del fatto che siano entrambe cautelative nei confronti dell'incertezza relativa al p -quantile, per la dimensione finita del campione di dati esaminati, il modo in cui quantificano questa incertezza è radicalmente diverso l'uno dall'altro.

Anche nel caso frequentista, infatti, la *design curve* è il luogo dei limiti inferiori di tollerabilità, $lti(y_x^p)$, del p -quantile:

$$lti(y_x^p) = \hat{\alpha} - \hat{\beta} \cdot x - \sqrt{\frac{1}{n} + \frac{x^2}{\sum_{i=1}^n x_i^2}} t_{\delta; 1-\gamma; n-2} \cdot s$$

ove $(1-\gamma)$ è il livello di confidenza e $t_{\eta; 1-\gamma; n-2}$ è l' $(1-\gamma)$ -esimo percentile della distribuzione di Student non-centrale, con $n-2$ gradi di libertà e parametro di non centralità $\eta = -z_p / \sqrt{1/n + x^2 / \sum_{i=1}^n x_i^2}$. Però l'incertezza è misurata, in questo caso, dal livello di confidenza, ovvero dalla probabilità che nel campionamento ripetuto l'intervallo definito dal limite di tollerabilità contenga il p -quantile stesso. Si tratta, quindi, di una misura iniziale di precisione, per la quale non potrà mai aversi sufficiente evidenza sperimentale.

Nel caso bayesiano, invece, il livello di credibilità con il quale ci si cautea dall'incertezza è una misura finale, *a posteriori*, costruita sulla base di ciò che è ben noto e consolidato con dati sperimentali.

$$D = 2^{(n-2)/2} \Gamma(n/2) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[\sum_{i=1}^n (y_i - \alpha + \beta x_i)^2 \right]^{-n/2} \pi(\alpha, \beta | UTS) d\alpha d\beta \quad (8)$$

$$\pi_2(\alpha, \beta | \Delta, UTS) = \frac{2^{(n-2)/2} \Gamma(n/2)}{D} \cdot \left[(n-2) \cdot s^2 + n \cdot (\alpha - \hat{\alpha})^2 + \left(\sum_{i=1}^n x_i^2 \right) \cdot (\beta - \hat{\beta})^2 \right]^{-\frac{n}{2}} \cdot \pi(\alpha, \beta | UTS) \quad (9)$$

5. CONFRONTO

Per verificare l'utilità e l'efficacia dell'informazione *a priori* nel migliorare l'accuratezza delle stime rispetto alla metodologia frequentista, sono state effettuate alcune simulazioni con il metodo Montecarlo. Tale metodo, com'è noto, consente di generare un numero grande a sufficienza di campioni di durate, per poter valutare le stime bayesiane di α e β , nonché i limiti di tollerabilità bayesiani, ed ottenere, infine, un campione di determinazioni degli stimatori bayesiani dal quale dedurre le loro proprietà.

Di questa analisi si riportano, per brevità, solo i primi risultati, relativi al caso di un acciaio virtuale ottenuto per laminazione a caldo, avente $UTS=1000$ MPa e parametri di fatica S'_f e b calcolati con le rette di regressione delle Fig. 1 e 2. Poiché, inoltre, la gran mole di dati sperimentali finora prodotta nel caso degli acciai consente di fissare per il coefficiente di variazione, cv , delle durate l'intervallo di variabilità (0.3 ÷ 0.5) [2], al fine di testare il comportamento degli stimatori bayesiani su campioni di dati affetti da un livello di scatter realistico, si è scelto di generare le durate da una popolazione log-normale avente c.v. pari al valore mediano di detto intervallo, al quale corrisponde una deviazione standard delle durate trasformate uguale a 0.167.

I campioni di durate sono stati ottenuti simulando, con una procedura descritta altrove [16], n. 2 prove per ciascuno dei tre valori di ampiezza, S_a , del carico: 0.35 UTS , 0.375 UTS e 0.4 UTS . Nella Tab. 2 sono riportate le stime delle errori quadratici medi (eqm) e dei bias

relativi degli stimatori bayesiani $\hat{\alpha}_B$ e $\hat{\beta}_B$, calcolati con un campione di 2000 determinazioni: si riconosce che entrambi gli stimatori bayesiani sono praticamente non distorti, risultando trascurabili i loro bias relativi. Inoltre, lo stimatore bayesiano di β risulta affetto da un errore quadratico medio molto minore dello stimatore frequentista $\hat{\beta}$, mentre le stime bayesiane e frequentiste di α risultano disperse quasi in ugual misura.

Con i dati di Tab. 3, anch'essi dedotti da un campione di 2000 simulazioni, è possibile eseguire inoltre un confronto tra *design curves*, frequentista e bayesiana. In essa sono riportati i valori medi dei limiti inferiori di tollerabilità frequentisti con livello di confidenza pari a 0.9 e degli analoghi limiti bayesiani, relativi al quantile $R = 0.95$ dei tre livelli di carico delle prove simulate, nonché le riduzioni percentuali delle distanze medie (r.d.) che separano le design curve, riduzioni realizzabili quando si passa dall'analisi frequentista e/o non informativa a quella bayesiana informativa. Come si vede, queste riduzioni sono minime ($\approx 10\%$) in corrispondenza del livello di carico intermedio e molto più significative ai livelli massimo e minimo, evidenziando una *design curve* bayesiana media più prossima, per forma e posizione, alla retta dei p -quantili del piano trasformato. Nonostante ciò, il suo grado di conservatività risulta essere pari a quello della curva frequentista, dal momento che le percentuali di coprimento (p.c.) dei limiti bayesiani, ovvero il numero di volte espresso in percentuale che nel campionamento ripetuto il limite è maggiore del corrispondente p -quantile, sono uguali al livello di confidenza dei limiti frequentisti.

	bias relativo	eqm
$\hat{\alpha}$	0	$4,67 \cdot 10^{-3}$
$\hat{\alpha}_B$	$4,93 \cdot 10^{-4}$	$4,40 \cdot 10^{-3}$
$\hat{\beta}$	0	8,321
$\hat{\beta}_B$	$1,18 \cdot 10^{-4}$	$8,98 \cdot 10^{-2}$

Tabella 2: Confronto tra stime puntuali frequentiste e bayesiane.

S_a	y_x^p	valori medi		r.d.	p.c.
		freq	bayes		
0,35 S_u	6,126	5,835	5,897	21,31 %	90,5%
0,375 S_u	5,773	5,519	5,550	12,21 %	89,8%
0,4 S_u	5,444	5,155	5,221	23,04 %	90,3%

Tabella 3: Confronto tra limiti inferiori di tollerabilità frequentisti e bayesiani.

6. CONCLUSIONI

Finora l'inferenza bayesiana è stata utilizzata per l'analisi dei dati di fatica di materiali solo con funzioni di densità a priori non informative. In questa nota, la metodologia è stata sfruttata per l'analisi di campioni di durate di piccola dimensione, con una funzione di densità a priori informativa per alcuni parametri del modello probabilistico che descrivere il legame carichi-durate. Ne è derivata una procedura di stima delle proprietà di fatica di implementazione numerica relativamente semplice, dal momento che richiede solo alcune quadrature di ordine non maggiore di tre.

L'efficacia delle funzioni a priori nell'arricchire le informazioni del campione di dati sperimentali è stata verificata, mediante uno studio pseudo-sperimentale, per un acciaio virtuale le cui proprietà di fatica ricadono nella zona di massima densità di probabilità *a priori* e con un grado di dispersione delle durate simile a quello degli acciai reali. I risultati ottenuti evidenziano che lo stimatore bayesiano della pendenza della retta delle durate trasformate medie è molto più efficiente di quello frequentista ed è affetto da un bias relativo trascurabile.

L'informazione a priori, infine, è in grado di condurre ad una *design curve* che si dimostra essere più realistica di quella frequentista, particolarmente in prossimità degli estremi del campo di valori di carico utilizzati per le prove, pur conservando lo stesso grado di prudenzialità.

7. BIBLIOGRAFIA

- [1] R.E. Little, in *Statistical Analysis of Fatigue Data* ASTM STP 744, R.E. Little and J.C. Ekvall Eds., (1981) 3-23.
 [2] P.H. Wirsching, in *Probabilistic Structural Mechanics Handbook – Theory and Industrial*

Applications, ed. C. (Raj) Sundararajan, Chapman & Hall, (1995).

- [3] J.F. Lawless J.F., *Statistical Models and Methods for Lifetime Data*, Wiley, New York, USA (1982).
 [4] C.L. Shen, P.H. Wirsching, G.T. Cashman, *Journal of Engineering Materials and Technology*, 118 (1996) 535-541.
 [5] G.E.P. Box, G.C. Tiao, *Bayesian Inference in Statistical Analysis*, Addison-Wesley Pub. Co, Reading (MA), USA (1973).
 [6] M.Akama, *Reliability Engineering & System Safety*, 75-3 (2002) 321-332.
 [7] R. Zhang, S. Mahadevan, *Structural Safety*, 22-2 (2000) 145-160.
 [8] E.H. Zavoni, R.M. Iturrizaga, *Journal of Offshore Mechanics and Arctic Engineering*, 126-3 (2004) 243-249.
 [9] H.O. Madsen, in *Probabilistic Methods in the Mechanics of Solids and Structures* S. Eddertz and N.C. Lind Eds., Proc. IUTAM Symp., Stockholm, Sweden, (1984)395-406.
 [10] G. Edwards, L.A. Pacheco, *Structural Safety*, 2 (1984) 27-38.
 [11] M.L. Roessle, A. Fatemi, *Int. J. Fatigue*, 22 (2000) 495-511.
 [12] M.A. Meggiolaro, J.T.P. Castro, *Int. J. Fatigue*, 26 (2004) 463-476.
 [13] J.H. Ong, *Int. J. Fatigue*, 15-1 (1993) 13-19.
 [14] J.H. Park, J.H. Song, *Int. J. Fatigue*, 24 (2002) 783-793.
 [15] M.R. Mitchell, in *ASM Handbook, Fatigue and Fracture*, Lampman S. ed., ASM International, 19 (1996).
 [16] A. De Iorio, M. Guida, F. Penta, P. Pinto, *Atti del XXXVII Convegno Nazionale AIAS, Associazione Nazionale Analisi delle Sollecitazioni*, Università di Roma "La Sapienza" (sett. 2008).

8. APPENDICE

$$\begin{aligned} \text{Poichè } \sum_{i=1}^n x_i &= \sum_{i=1}^n (u_i - \bar{u}) = \sum_{i=1}^n u_i - n \cdot \bar{u} = 0 \\ \sum_{i=1}^n (y_i - \hat{a} + \hat{b}x_i) &= \sum_{i=1}^n y_i - n \cdot \hat{a} + \hat{b} \cdot \sum_{i=1}^n x_i = 0 \\ \sum_{i=1}^n (y_i - \hat{a} + \hat{b}x_i) \cdot x_i &= \sum_{i=1}^n x_i y_i - \hat{a} \cdot \sum_{i=1}^n x_i + \hat{b} \cdot \sum_{i=1}^n x_i^2 = 0 \end{aligned}$$

sommando e sottraendo in ciascun termine della sommatoria $\sum_{i=1}^n (y_i - a + bx_i)^2$ la quantità $\hat{a} - \hat{b} \cdot x_i$, si ottiene

$$\begin{aligned} \sum_{i=1}^n (y_i - a + bx_i)^2 &= \sum_{i=1}^n (y_i - \hat{a} + \hat{b}x_i + \hat{a} - \hat{b}x_i - a + bx_i)^2 = \\ &= \sum_{i=1}^n \left[(y_i - \hat{a} + \hat{b}x_i) - (a - \hat{a}) + x_i \cdot (b - \hat{b}) \right]^2 = \\ &= (n-2) \cdot s^2 + n \cdot (a - \hat{a})^2 + (b - \hat{b})^2 \cdot \sum_{i=1}^n x_i^2 \end{aligned}$$