

# POLITECNICO DI TORINO

## Corso di Laurea Magistrale in Ingegneria Matematica



### Valutazione dell'incertezza di misura nei Digital Twin con Statistica Bayesiana

#### Relatori

Professor Gianluca Mastrantonio  
Professor Giacomo Maculotti

#### Candidato

Geard Koci

Anno Accademico 2024-2025



# Introduzione

La caratterizzazione meccanica dei materiali rappresenta un aspetto fondamentale nel controllo dei processi produttivi e nella valutazione delle prestazioni dei prodotti, particolarmente rilevante nel contesto dell’Industria 4.0 e delle iniziative di sostenibilità come il Green Deal europeo [30].

Tra le tecniche di caratterizzazione meccanica, l’Instrumented Indentation Test (IIT) o nanoindentazione si distingue come metodo altamente informativo, in grado di fornire proprietà meccaniche quali durezza, modulo elastico e tensioni residuali attraverso misurazioni non distruttive [35].

La nanoindentazione consiste nell’applicazione di una forza controllata mediante un indentatore su un campione, misurando simultaneamente la forza applicata  $F$  e la profondità di penetrazione  $h$  durante l’intero ciclo di prova.

Il comportamento del materiale durante l’indentazione può essere descritto matematicamente attraverso la relazione di Boussinesq:

$$F = a(h - h_0)^m \quad (1)$$

dove  $a$  e  $m$  sono parametri che dipendono dal materiale e dalla geometria dell’indentatore, mentre  $h_0$  rappresenta il punto di primo contatto durante la fase di carico o l’indentazione plastica residua  $h_p$  durante la fase di scarico.

La piattaforma utilizzata per le misurazioni è un’Anton Paar STeP6 dotata di testa NHT3, disponibile presso la sala metrologica del laboratorio Mind4Lab (DIGEP – Politecnico di Torino) [36]. Il sistema è equipaggiato con un indentatore Berkovich in diamante drogato al boro, correttamente calibrato, con test di riferimento condotti su silice fusa NPL in modalità a controllo di forza.

A partire dai dati forniti dal DIGEP, questa tesi si propone di analizzare le misurazioni ottenute in condizioni ambientali controllate e non controllate, stimando i parametri  $h_0$ ,  $a$  e  $m$  della relazione di Boussinesq in modo robusto e interpretabile.

Dopo una fase iniziale di visualizzazione ed esplorazione dei dataset, l’attività si concentrerà sulla stima dei parametri del modello mediante tecniche di regressione. In particolare, verrà adottato un approccio bayesiano per ottenere stime più affidabili e quantificare l’incertezza associata ai parametri, con l’obiettivo finale di migliorare la rappresentazione del Digital Twin del processo di indentazione.

I risultati ottenuti contribuiranno allo sviluppo di una modellazione più solida, utile per ottimizzare le procedure di misura e abilitare applicazioni avanzate nella caratterizzazione meccanica dei materiali, in linea con gli obiettivi di zero-defect manufacturing e sostenibilità industriale.

Per far ciò utilizzeremo la libreria STAN per determinare in maniera statisticamente accurata i parametri, confrontando anche diversi modelli fra di loro, attraverso grafici e parametri di convergenza.

L'obiettivo finale sarà quello poi di presentare due notebook python distinti, che sfruttano il modello trovato nel punto precedente, per generare sinteticamente nuovi dataset, oppure per capire se, quando ottenuto un nuovo punto in maniera sperimentale, esso segua l'andamento teorico che abbiamo previsto.

# Indice

<b>1</b>	<b>Nanoindentazione e Digital Twins</b>	1
1.1	Introduzione alla nanoindentazione . . . . .	1
1.1.1	Il Test di Indentazione Strumentale . . . . .	1
1.1.2	Applicazioni e Sfide della Nanoindentazione . . . . .	2
1.1.3	Approccio Metodologico . . . . .	2
1.2	Digital Twins e le loro applicazioni . . . . .	3
1.2.1	Architettura dei Digital Twins . . . . .	3
1.2.2	Applicazioni nei sistemi di produzione . . . . .	4
1.2.3	Digital Twins in metrologia . . . . .	5
1.2.4	Il caso specifico: Digital Twin per la nanoindentazione	5
1.2.5	Vantaggi e prospettive future . . . . .	7
1.3	Introduzione alla statistica Bayesiana . . . . .	7
1.3.1	Scelta del prior e verosimiglianza . . . . .	8
1.3.2	Metodi computazionali . . . . .	8
1.3.3	Applicazioni . . . . .	9
<b>2</b>	<b>Analisi Esplorativa e Modellazione Preliminare</b>	10
2.1	Esplorazione e Visualizzazione Dati . . . . .	10
2.1.1	Descrizione del Dataset Sperimentale . . . . .	10
2.1.2	Struttura dei Dati e Parametri di Configurazione .	11
2.1.3	Metodologia di Visualizzazione e Analisi Preliminare	11
2.1.4	Analisi del Dataset 1: Condizioni Senza Controllo Termico . . . . .	12
2.1.5	Analisi del Dataset 2: Condizioni con Controllo Termico . . . . .	13
2.1.6	Analisi Comparativa tra i due Dataset . . . . .	13
2.1.7	Interpretazione Fisica dei Risultati . . . . .	14
2.2	Calcolo dei Parametri del Modello mediante Verosimiglianza	15
2.2.1	Sviluppo del Modello Matematico . . . . .	15
2.2.2	Funzione di Verosimiglianza . . . . .	16
2.2.3	Strategia di Ottimizzazione: Differential Evolution .	17

2.2.4	Definizione dei Vincoli e Inizializzazione . . . . .	18
2.2.5	Aggregazione Multi-Esperimento . . . . .	18
2.2.6	Validazione del Modello . . . . .	19
2.3	Risultati Sperimentali . . . . .	19
2.3.1	Confronto Grafico del Fitting per il Dataset 1 e 2 .	20
2.3.2	Confronto dei Parametri tra Dataset . . . . .	22
2.3.3	Qualità del Fitting e Precisione delle Stime . . . . .	24
2.3.4	Stabilità delle Fasi di Mantenimento . . . . .	24
2.3.5	Implicazioni per la Caratterizzazione Meccanica . .	24
<b>3</b>	<b>Analisi Bayesiana con Stan per la Nanoindentazione</b>	<b>26</b>
3.1	Introduzione a Stan e all'Analisi Bayesiana . . . . .	26
3.1.1	Il Framework Bayesiano . . . . .	26
3.1.2	Algoritmo Hamiltonian Monte Carlo (HMC) . . . .	27
3.2	Preparazione e Caricamento dei Dati . . . . .	27
3.3	Implementazione in Stan . . . . .	28
3.3.1	Struttura del Modello Stan . . . . .	28
3.3.2	Strategie di Inizializzazione e Parametri . . . . .	31
3.4	Diagnostiche di Convergenza . . . . .	32
3.4.1	R-hat Statistic ( $\hat{R}$ ) . . . . .	32
3.4.2	Effective Sample Size (ESS) . . . . .	33
3.5	Risultati del Modello . . . . .	34
3.5.1	Diagnostiche per Parametro - Dataset 1 . . . . .	34
3.5.2	Diagnostiche per Parametro - Dataset 2 . . . . .	35
3.5.3	Stime a Posteriori dei Parametri - Dataset 1 . . . .	36
3.5.4	Stime a Posteriori dei Parametri - Dataset 2 . . . .	36
3.5.5	Confronto degli Intervalli di Credibilità tra Dataset	37
3.5.6	Criteri di Validazione tramite RMSE . . . . .	39
3.5.7	Metriche di Performance - Dataset 1 . . . . .	39
3.5.8	Metriche di Performance - Dataset 2 . . . . .	40
3.6	Modelli Alternativi: Log-Normale e Gamma . . . . .	42
3.6.1	Modello Log-Normale . . . . .	42
3.6.2	Modello Gamma . . . . .	43
3.7	Risultati dei Modelli Alternativi . . . . .	44
3.7.1	Modello Log-Normale - Dataset 1 . . . . .	44
3.7.2	Modello Log-Normale - Dataset 2 . . . . .	45
3.7.3	Confronto degli Intervalli di Credibilità - Modello Log-Normale . . . . .	49
3.7.4	Modello Gamma - Dataset 1 . . . . .	51
3.7.5	Modello Gamma - Dataset 2 . . . . .	52

3.7.6	Confronto degli Intervalli di Credibilità - Modello Gamma . . . . .	55
3.7.7	Sintesi Comparativa Multi-Modello . . . . .	56
3.8	Analisi Comparativa dei Risultati e Conclusioni . . . . .	57
3.8.1	Sintesi dei Risultati Ottenuti . . . . .	57
3.8.2	Il Problema dell’Ottimizzazione Locale . . . . .	58
3.8.3	Superiorità del Modello Normale: Evidenze Teoriche e Pratiche . . . . .	58
3.8.4	Raccomandazioni Metodologiche . . . . .	59
3.8.5	Limitazioni e Sviluppi Futuri . . . . .	60
3.8.6	Conclusioni Finali . . . . .	60
<b>4</b>	<b>Riproduzione del modello fisico</b>	<b>62</b>
4.1	Implementazione del Modello a Cinque Fasi . . . . .	62
4.1.1	Scelta dei Parametri del Modello . . . . .	62
4.1.2	Definizione del Profilo di Carico Temporale . . . . .	63
4.1.3	Modello Matematico per il Calcolo degli Spostamenti	64
4.1.4	Garanzia di Continuità del Modello . . . . .	64
4.1.5	Validazione e Output . . . . .	65
4.1.6	Dataset senza controllo termico . . . . .	67
4.1.7	Dataset con controllo termico . . . . .	67
4.2	Validazione dei Dati Sperimentali . . . . .	69
4.2.1	Approccio Metodologico . . . . .	69
4.2.2	Struttura dei Dati Sperimentali . . . . .	70
4.2.3	Calcolo degli Spostamenti Teorici . . . . .	70
4.2.4	Metriche di Validazione . . . . .	70
4.2.5	Gestione dei Parametri del Modello . . . . .	72
4.2.6	Output e Visualizzazione . . . . .	72
4.2.7	Dataset senza controllo termico . . . . .	72
4.2.8	Dataset con controllo termico . . . . .	75
4.2.9	Dataset teorico e sperimentale diversi . . . . .	76
<b>Bibliografia</b>		<b>79</b>



# Capitolo 1

# Nanoindentazione e Digital Twins

In questo capitolo esploreremo innanzitutto la tecnica della nanoindentazione, illustrando i principi sperimentali e le informazioni meccaniche che essa fornisce a scala nanometrica; successivamente, introdurremo il concetto di Digital Twin, descrivendo come queste repliche virtuali consentano simulazioni predittive e monitoraggio in tempo reale di sistemi fisici; continuando, presenteremo le basi della statistica Bayesiana, evidenziando il ruolo del teorema di Bayes nella formulazione di inferenze probabilistiche e i principali metodi computazionali per l'analisi dei dati; infine, confronteremo le strategie di modellazione tradizionali con quelle bayesiane, mettendo in luce vantaggi, limiti e possibili integrazioni tra approcci deterministicici e probabilistici.

## 1.1 Introduzione alla nanoindentazione

### 1.1.1 Il Test di Indentazione Strumentale

Il Test di Indentazione Strumentale (Instrumented Indentation Test, IIT) rappresenta una tecnica di caratterizzazione meccanica altamente informativa, in grado di fornire proprietà fondamentali come durezza, modulo elastico e tensione residua [35].

Questa tecnica consiste nell'applicare una forza mediante un indentatore su un campione da caratterizzare, misurando simultaneamente la forza applicata  $F$  e la profondità di penetrazione  $h$  nel materiale durante l'intera durata del test [44].

L'analisi della porzione di scarico della curva di indentazione (IC), ovvero  $F(h)$ , permette la valutazione delle proprietà meccaniche attraverso la relazione:

$$h_{c,max} = h_{max} - h_0 - C_f F_{max} - \varepsilon \left( \frac{1}{S_m} - C_f \right) F_{max} \quad (1.1)$$

dove  $h_0$  è il punto di primo contatto,  $C_f$  è la compliance del telaio,  $\varepsilon$  è il parametro geometrico relativo all'indentatore,  $S_m$  è la rigidezza di contatto, e  $F_{max}$  è la forza massima applicata.

Da queste misurazioni si possono derivare la durezza di indentazione  $H_{IT}$  e il modulo di indentazione  $E_{IT}$ :

$$H_{IT} = \frac{F_{max}}{A_p(h_{c,max})} \quad (1.2)$$

$$E_{IT} = \frac{1 - \nu_s^2}{\frac{2}{\sqrt{A_p(h_{c,max})}} \frac{S}{\sqrt{\pi}} - \frac{1 - \nu_i^2}{E_i}} \quad (1.3)$$

dove  $A_p(h_{c,max})$  è la funzione di area dell'indentatore,  $\nu_s$  e  $\nu_i$  sono i rapporti di Poisson del campione e dell'indentatore, ed  $E_i$  è il modulo di Young dell'indentatore [43].

### 1.1.2 Applicazioni e Sfide della Nanoindentazione

La nanoindentazione trova ampie applicazioni nei settori aerospaziale, automobilistico, dei semiconduttori, nucleare e biomedicale per ottimizzare i processi produttivi, caratterizzare microstrutture e valutare le proprietà meccaniche di materiali multistrato [30]. Particolare rilevanza assume nella caratterizzazione di rivestimenti ultrasottili (inferiori a 100 nm) per applicazioni ottiche ed elettroniche, settore che rappresenta un mercato globale di 50 miliardi di dollari USA [37].

Tuttavia, la tecnica IIT è altamente sensibile a diverse condizioni sperimentali e può essere influenzata da numerosi fattori di influenza. Oltre alla già menzionata compliance del telaio e geometria dell'indentatore, le condizioni del campione e ambientali giocano un ruolo cruciale nel determinare l'accuratezza e la precisione dei risultati di caratterizzazione. Alcuni fattori possono causare errori casuali (topografia superficiale, vibrazioni meccaniche e rumore), altri errori sistematici (smussamento della punta dell'indentatore, campo termico non omogeneo al contatto), mentre altri ancora inducono derive (usura progressiva dell'indentatore dovuta ad attrito e affinità chimica).

### 1.1.3 Approccio Metodologico

Per affrontare queste sfide, il lavoro presenta una metodologia per stabilire un Digital Twin (DT) tracciabile per la nanoindentazione. Per la

realizzazione del DT tracciabile, viene adottato un approccio data-driven informato dalla fisica. Conoscendo le relazioni teoriche che descrivono la curva di indentazione secondo il problema di Boussinesq:

$$F = a(h - h_0)^m \quad (1.4)$$

dove  $a$  e  $m$  dipendono dal materiale del campione e dalla geometria dell'indentatore, viene utilizzata una regressione non lineare per stimare i parametri del modello  $\vartheta = \{a, m, h_0\}$ .

Il caso studio presentato utilizza una piattaforma Anton Paar STeP6 equipaggiata con una testina di nanoindentazione NHT3 e un indentatore Berkovich in diamante drogato con boro. Vengono eseguite 5 misurazioni replicate su silice fusa (materiale di riferimento NPL calibrato) in controllo di forza, con fasi di carico, mantenimento e scarico di 30 s, 10 s e 30 s rispettivamente.

La tracciabilità del DT viene stabilita attraverso la calibrazione dei trasduttori forza-spostamento, della compliance del telaio e della geometria dell'indentatore.

## 1.2 Digital Twins e le loro applicazioni

Un Digital Twin (DT) è una rappresentazione virtuale di un oggetto o sistema progettata per riflettere accuratamente un oggetto fisico. Più precisamente, un digital twin è definito come 'un modello di simulazione che replica accuratamente un sistema fisico in un ambiente virtuale e che include aggiornamenti dinamici automatizzati del modello virtuale secondo lo stato osservato della sua controparte reale per ottenere un controllo fisico il più automatizzato possibile di quest'ultima'. [11].

Il concetto di Digital Twin si basa su una connessione dinamica bidirezionale tra il mondo fisico e quello virtuale. I digital twins sono collegati a fonti di dati reali dall'ambiente, il che significa che il twin si aggiorna in tempo reale, permettendo ai produttori di monitorare le prestazioni, prevedere problemi e ottimizzare i processi basandosi su dati live [11, 28].

### 1.2.1 Architettura dei Digital Twins

Secondo la letteratura classica, un Digital Twin presenta quattro elementi distintivi [21]:

1. **Ambiente fisico:** include le entità fisiche a cui il DT fa riferimento
2. **Ambiente virtuale:** comprende le entità virtuali che modellano le entità fisiche considerate

3. **Connessione Physical-to-Virtual (P2V)**: stabilisce il collegamento tra l'ambiente fisico e quello virtuale
4. **Connessione Virtual-to-Physical (V2P)**: implementa strategie di controllo per raggiungere le prestazioni target del sistema fisico

Questa architettura bidirezionale consente non solo il monitoraggio passivo ma anche il controllo attivo dei sistemi fisici, stabilendo un flusso di dati continuo e dinamico.

Per un DT di uno strumento di misura, l'entità virtuale può essere considerata un esperimento virtuale variabile nel tempo:

$$X = g(Y, W, t; \vartheta) \quad (1.5)$$

dove  $X$  rappresenta le osservazioni (misure),  $Y$  il misurando (sconosciuto),  $W$  i parametri di misura e le condizioni al contorno,  $t$  il tempo e  $\vartheta$  i parametri del modello.

### 1.2.2 Applicazioni nei sistemi di produzione

I digital twins nella produzione creano repliche virtuali di processi produttivi, attrezzature e intere operazioni di fabbrica. Le principali applicazioni includono [28, 38]:

- **Monitoraggio e controllo**: Un digital twin può aiutare a monitorare lo stato, rilevare anomalie, prevedere comportamenti del sistema e prescrivere operazioni future
- **Manutenzione predittiva**: analisi della salute delle macchine e programmazione della manutenzione
- **Ottimizzazione dei processi**: valutazione di piani e programmi alternativi
- **Commissioning virtuale**: test e validazione di sistemi prima dell'implementazione fisica

I digital twins sono massicciamente sfruttati in letteratura per i processi di produzione, i processi di assemblaggio, il controllo e la previsione in tempo reale dello stato di salute di sistemi complessi, come edifici e aeromobili [56, 32].

### 1.2.3 Digital Twins in metrologia

Nel campo della metrologia, i Digital Twins assumono un ruolo particolarmente importante. I DT sono essenziali per correggere e compensare gli errori sistematici, implementare strategie di controllo finalizzate alla manutenzione predittiva, prevenire errori gravi e malfunzionamenti che potrebbero danneggiare l'apparato di misura e, infine, identificare le fonti di errori sistematici [23].

I vantaggi specifici nella metrologia includono [57]:

- **Simulazione offline:** quando i metrologi programmano i twin dei loro ambienti e attrezzature reali, il CMM o il robot sono liberi di continuare a misurare e monitorare
- **Prevenzione errori:** i programmi di ispezione simulati e testati virtualmente sono privi di errori e collisioni una volta applicati al processo produttivo
- **Interpretazione automatica:** capacità di lavorare direttamente con file CAD nativi o neutri e interpretare automaticamente GDT.

La disponibilità di DT metrologici, cioè tracciabili, di strumenti e processi di misura, permetterebbe di introdurre l'incertezza di misura in DT più complessi di processi e sistemi.

### 1.2.4 Il caso specifico: Digital Twin per la nanoindentazione

Il presente lavoro presenta un caso di studio innovativo per l'applicazione dei Digital Twins alla nanoindentazione, una tecnica di caratterizzazione meccanica ad alta risoluzione. La nanoindentazione è particolarmente sensibile a diversi fattori di influenza che possono introdurre errori sistematici o aumentare l'incertezza di misura.

Il Digital Twin metrologico proposto per la nanoindentazione si basa su una strategia di modellazione data-driven informata dalla fisica [31], che prevede:

#### Modellazione Physical-to-Virtual (P2V)

La modellazione P2V utilizza l'equazione teorica di Boussinesq per l'indentazione [34]:

$$F = a(h - h_0)^m \quad (1.6)$$

dove  $F$  è la forza applicata,  $h$  è la profondità di penetrazione,  $a$  e  $m$  dipendono dal materiale del campione e dalla geometria dell'indentatore,

e  $h_0$  è il punto di primo contatto durante il carico o la deformazione plastica residua durante lo scarico.

Questa strategia permette di mantenere la fisica del problema pur sfruttando dati empirici per la calibrazione.

Il sistema è progettato per identificare automaticamente:

- Errori di misura e derive temporali
- Fattori sistematici che influenzano le prestazioni
- Anomalie nel comportamento dell'indentatore

Le quantità di controllo monitorate includono la curva di indentazione, i risultati della caratterizzazione meccanica (durezza  $H_{IT}$  e modulo  $E_{IT}$ ), quantità rilevanti che descrivono il comportamento del modello di indentazione (rigidezza di contatto  $S$ ), e possibili quantità aggiuntive misurate come temperatura  $T$  e resistenza di contatto elettrico  $R$ .

Il sistema inoltre implementa strategie di controllo che permettono:

- Correzione automatica delle misure
- Manutenzione predittiva del sistema
- Identificazione delle cause principali degli errori attraverso Root Cause Analysis (RCA)

Se viene identificata un'indentazione fuori controllo, la causa principale dovrebbe essere rintracciata e le strategie di controllo pertinenti implementate per correggere, compensare o eliminare la fonte dell'errore.

Un aspetto cruciale del Digital Twin proposto è la sua tracciabilità metropolitana, ottenuta attraverso:

- Calibrazione dei trasduttori forza-spostamento
- Calibrazione della compliance del telaio e della geometria dell'indentatore [36]
- Valutazione dell'incertezza mediante approcci simulativi non parametrici (Bootstrap)

L'incertezza di misura è essenziale per identificare le condizioni fuori controllo con livello di confidenza statistica. La valutazione dell'intervallo di predizione del modello dell'entità virtuale  $g$  permette di controllare la dispersione dello spostamento e propagare i contributi rilevanti alla caratterizzazione meccanica [37].

### 1.2.5 Vantaggi e prospettive future

La tecnologia digital twin aiuta le organizzazioni a utilizzare dati in tempo reale, mockup, algoritmi e AI per creare una rappresentazione virtuale di un oggetto fisico, processo o servizio. Tali repliche sono tipicamente più convenienti dei prototipi fisici, permettendo alle organizzazioni di creare e testare versioni illimitate [47].

Per il caso specifico della nanoindentazione, il Digital Twin metrologico offre:

1. **Identificazione automatica di errori:** capacità di rilevare misure fuori controllo con confidenza statistica
2. **Manutenzione predittiva:** prevenzione di malfunzionamenti gravi che potrebbero danneggiare l'apparato di misura
3. **Controllo qualità industriale:** applicabilità su larga scala per il controllo statistico di qualità
4. **Riduzione barriere d'ingresso:** minore necessità di esperienza specifica per distinguere artefatti di misura da fenomeni reali del materiale

Il futuro sviluppo si concentrerà sull'addestramento di algoritmi di intelligenza artificiale per identificare automaticamente le cause principali degli errori e implementare di conseguenza strategie di correzione, portando verso una completa automazione del processo di nanoindentazione.

## 1.3 Introduzione alla statistica Bayesiana

La statistica Bayesiana è un paradigma di inferenza che interpreta la probabilità come grado di credenza soggettiva piuttosto che frequenza relativa. Il fulcro teorico è il *teorema di Bayes*, che trasforma le conoscenze pregresse su un parametro sconosciuto  $\theta$  nella distribuzione *a posteriori* una volta osservati i dati  $D$ :

$$p(\theta | D) = \frac{p(D | \theta) \pi(\theta)}{\int_{\Theta} p(D | \theta') \pi(\theta') d\theta'} \propto \underbrace{p(D | \theta)}_{\text{verosimiglianza}} \times \underbrace{\pi(\theta)}_{\text{prior}}.$$

Qui  $\pi(\theta)$  è la *distribuzione a priori* che incorpora informazioni ed esperienze pregresse, mentre

$$p(D | \theta) = \prod_{i=1}^n f(x_i | \theta)$$

è la *funzione di verosimiglianza*, basata sul modello di densità  $f(x | \theta)$  [8, 18].

### 1.3.1 Scelta del prior e verosimiglianza

La definizione del prior riveste un ruolo cruciale: priors informativi permettono di introdurre conoscenza specialistica, ma possono introdurre bias se mal calibrati [2]. In alternativa, priors non informativi (o debolmente informativi) cercano di avere impatto minimo sull'analisi. Tra i priors più utilizzati:

- *Jeffreys prior*, invariante sotto trasformazioni di parametro [27];
- *Priors coniugati*, che mantengono la forma della distribuzione a posteriori [2].

La funzione di verosimiglianza nasce dalla modellazione statistica dei dati (modelli parametrici o non parametrici): per il caso gaussiano, ad esempio, si usa

$$f(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

### 1.3.2 Metodi computazionali

Calcolare esplicitamente la normalizzazione nel teorema di Bayes è spesso irrealistico: si ricorre quindi a metodi numerici avanzati:

1. **Simulazione MCMC** (Metropolis–Hastings, Gibbs Sampling), che genera campioni dalla posteriori con garanzie di convergenza asintotica [40, 16, 20].
2. **Algoritmi di campionamento avanzato**: Hamiltonian Monte Carlo (HMC) e No-U-Turn Sampler (NUTS), implementati in Stan [7].
3. **Ottimizzazione euristica** per il punto MAP (Maximum A Posteriori), ad esempio Differential Evolution [50], Particle Swarm Optimization [33] o algoritmi genetici.
4. **Metodi a varianza ridotta** e approssimazioni deterministiche: Laplace Approximation, Expectation Propagation, Variational Bayes [29].

### 1.3.3 Applicazioni

- **Decision theory**: analisi di rischi e utilità, regole di decisione Bayesiane ottimali [1].
- **Machine Learning**: reti neurali Bayesiane, Gaussian Processes per regressione non lineare, modelli di misture infiniti [42, 46].
- **Economia e finanza**: stima di modelli GARCH Bayesiani, Value at Risk con incertezza parametrica [24].
- **Bioinformatica e medicina**: analisi di espressione genica, modelli per trial clinici adattativi [3].
- **A/B testing** in ambito digitale: decisioni in tempo reale basate su distribuzioni di probabilità predittive [5].

# Capitolo 2

## Analisi Esplorativa e Modellazione Preliminare

In questo capitolo viene presentata un'analisi iniziale dei dataset a disposizione, con l'obiettivo di evidenziarne le principali caratteristiche e le differenze tra i due insiemi di dati. L'esplorazione è condotta attraverso rappresentazioni grafiche, distinte per set e per singolo test. L'analisi si conclude con una prima stima dei parametri del modello, suddivisa per segmenti.

### 2.1 Esplorazione e Visualizzazione Dati

#### 2.1.1 Descrizione del Dataset Sperimentale

Il dataset utilizzato in questa analisi è composto da due insiemi distinti di dati sperimentali, acquisiti in condizioni differenziate. Il primo dataset è stato acquisito con l'aria condizionata della sala metrologica spenta, mentre il secondo con l'aria condizionata attiva. Questa distinzione metodologica è di fondamentale rilevanza per la qualità e l'affidabilità dei risultati sperimentali.

Durante il processo di nanoindentazione, la punta metallica dell'indattore può generare calore a causa dell'attrito e della deformazione plastica del materiale. Tale calore, accumulandosi nell'ambiente circostante, può influenzare significativamente la stabilità termica dello strumento di misura e, di conseguenza, la precisione delle misurazioni. Un controllo accurato della temperatura ambientale è quindi essenziale per garantire la ripetibilità dei risultati e minimizzare gli errori sistematici legati alle variazioni termiche.

Entrambi i dataset sono strutturati in modo identico e sono suddivisi in tre set di test distinti, ciascuno dei quali contiene 16 prove sperimentali

indipendenti, a meno del terzo set del secondo dataset che contiene 17 prove. Ogni singola prova è memorizzata in un file di testo dedicato, per un totale di 48 file per il Dataset 1 e 49 per il Dataset 2.

### 2.1.2 Struttura dei Dati e Parametri di Configurazione

I file di dati presentano una struttura standardizzata che facilita l’analisi sistematica. Ogni file contiene una sezione iniziale con i parametri di configurazione del test, tra cui il tempo di loading e unloading, la forza massima applicata, la velocità di carico e scarico, e altri parametri strumentali.

La parte centrale dei file contiene i dati effettivi della prova sperimentale, distribuiti su sei colonne principali: *Tempo (s)*, *Profondità (nm)*, *Forza (mN)*, *ForzaRef (mN)*, *SegmentID* e *Resistenza (ohm)*. Ogni prova registra approssimativamente 900 istanti temporali, ciascuno associato a valori di forza applicata e profondità raggiunta dall’indentatore.

Il campo *SegmentID* risulta particolarmente utile per l’analisi dei dati, in quanto permette di distinguere chiaramente tra le diverse fasi del test di nanoindentazione: la fase di carico (loading), la fase di mantenimento del carico (holding) e la fase di scarico (unloading). Questa classificazione facilita un’analisi più accurata del comportamento meccanico del materiale in ciascuna fase specifica del ciclo di indentazione.

### 2.1.3 Metodologia di Visualizzazione e Analisi Preliminare

Come primo passo dell’analisi esplorativa, vengono presentati i grafici delle grandezze medie di profondità e forza in funzione del tempo, calcolati per ogni set di dati. Questi grafici permettono di identificare immediatamente le caratteristiche temporali del processo di indentazione e di valutare la consistenza dei dati all’interno di ciascun set.

Successivamente, per ciascuno dei tre set viene riportato un grafico della profondità in funzione della forza, che rappresenta la classica curva di indentazione (load-displacement curve). Questi grafici sono fondamentali per osservare il comportamento caratteristico del materiale durante il ciclo di carico-scarico e per identificare eventuali fenomeni di isteresi o anomalie nel comportamento meccanico.

Tutti i grafici di questa prima parte sono stati realizzati utilizzando il linguaggio di programmazione Python, sfruttando le librerie matplotlib [26] e pandas [39] per la visualizzazione e l’elaborazione dei dati. I dati sono stati preliminarmente organizzati in strutture dataframe per facilitare le operazioni di calcolo statistico e la generazione dei grafici.

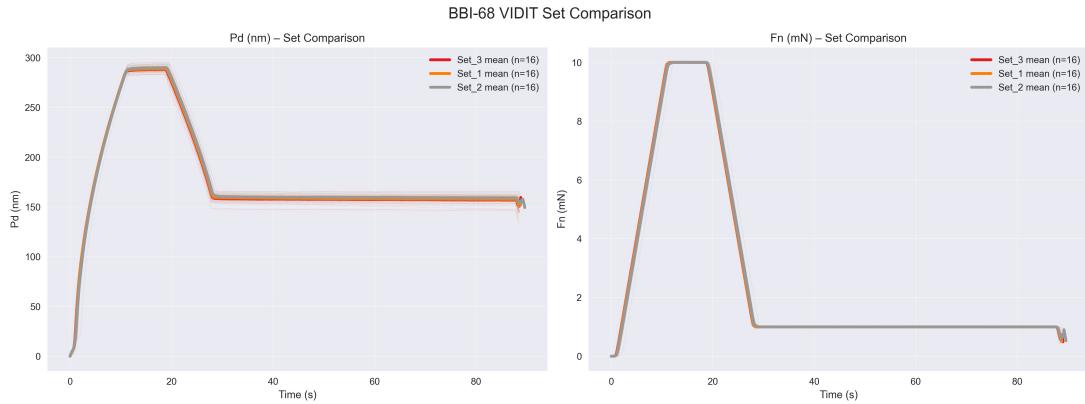


Figura 2.1. Profondità e forze medie nel tempo

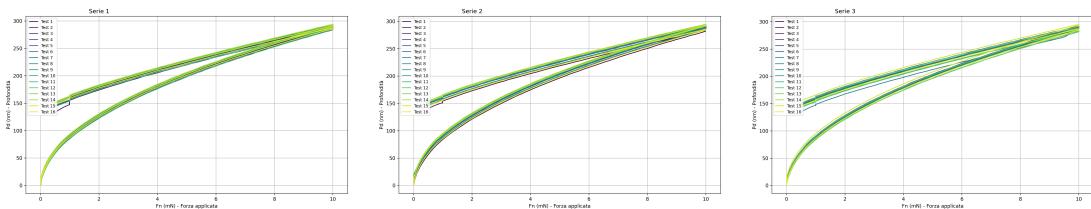


Figura 2.2. \*

Figura 2.3. \*

Figura 2.4. \*

Profondità e Forza - Set 1

Profondità e Forza - Set 2

Profondità e Forza - Set 3

Figura 2.5. Confronto tra i tre set per il dataset 1.

#### 2.1.4 Analisi del Dataset 1: Condizioni Senza Controllo Termico

I grafici del primo dataset mostrano le caratteristiche tipiche del processo di nanoindentazione in condizioni ambientali non controllate. Come previsto teoricamente, forza e profondità seguono la relazione di Boussinesq, di cui successivamente andremo a stimare i parametri.

I grafici temporali permettono di distinguere molto chiaramente i dati nelle diverse fasi operative: la fase di carico, caratterizzata da un aumento progressivo sia della forza che della profondità; la fase di mantenimento, in cui la forza rimane costante mentre la profondità può presentare piccole variazioni, e infine la fase di scarico, dove la forza diminuisce progressivamente mentre la profondità mostra il recupero elastico del materiale.

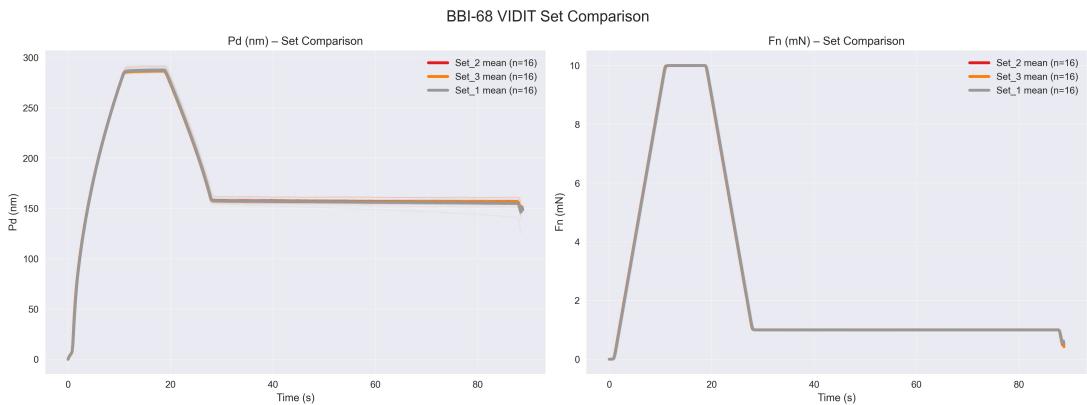


Figura 2.6. Profondità e forze medie nel tempo

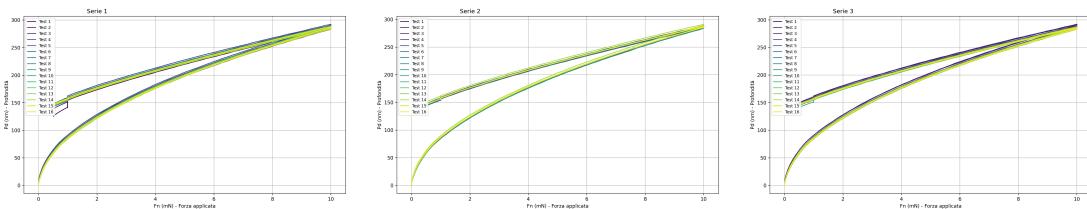


Figura 2.7. \*

Figura 2.8. \*

Figura 2.9. \*

Profondità e Forza - Set 1      Profondità e Forza - Set 2      Profondità e Forza - Set 3

Figura 2.10. Confronto tra i tre set per il dataset 2.

## 2.1.5 Analisi del Dataset 2: Condizioni con Controllo Termico

I risultati del secondo dataset, acquisito con l'aria condizionata attiva, vengono presentati di seguito per permettere un confronto diretto con le condizioni precedenti.

L'analisi visiva dei grafici del secondo dataset rivela caratteristiche molto simili a quelle del primo dataset in termini di forma generale delle curve e comportamento temporale. Tuttavia, le differenze più significative emergono dall'analisi comparativa quantitativa piuttosto che dall'osservazione diretta dei singoli grafici.

## 2.1.6 Analisi Comparativa tra i due Dataset

Per evidenziare le differenze tra i due dataset, che non risultano immediatamente evidenti dall'osservazione diretta dei grafici individuali, è stata condotta un'analisi comparativa quantitativa. Vengono presentati tre

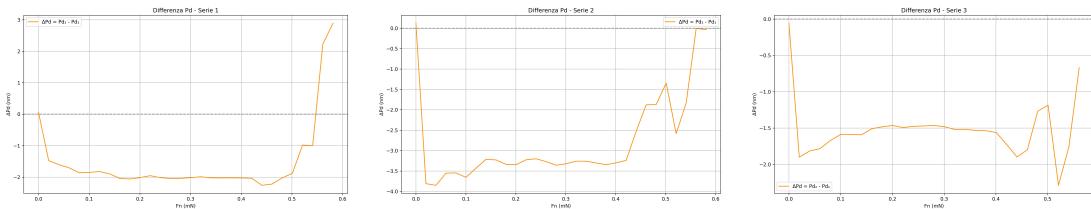


Figura 2.11. Set 1

Figura 2.12. Set 2

Figura 2.13. Set 3

Figura 2.14. Confronto tra la differenza nei tre set.

grafici, uno per ogni set, che mostrano la differenza di profondità tra il secondo e il primo dataset, a parità di forza applicata.

L’analisi è stata condotta calcolando la differenza punto per punto tra le curve di profondità dei due dataset, considerando un intervallo di confidenza ottenuto sottraendo e aggiungendo una deviazione standard ai valori medi. Questo approccio permette di quantificare l’incertezza associata alle misurazioni e di valutare la significatività statistica delle differenze osservate.

### 2.1.7 Interpretazione Fisica dei Risultati

L’analisi dei grafici differenziali rivela chiaramente che quando l’aria condizionata è attiva nella sala metrologica si ottengono profondità di indentazione sistematicamente minori rispetto alle condizioni senza controllo termico. Questo fenomeno ha una spiegazione fisica ben definita legata agli effetti della deriva termica (thermal drift) sul sistema di misura.

La presenza dell’aria condizionata durante la nanoindentazione limita efficacemente la deriva termica, ovvero l’espansione o contrazione del campione e del sistema strumentale dovuta a fluttuazioni termiche ambientali. La stabilizzazione termica si traduce in misurazioni più stabili e precise della profondità di indentazione, riducendo significativamente gli errori sistematici causati dall’accumulo di calore durante il test.

In condizioni termiche controllate, l’equilibrio termico mantenuto tra la punta dell’indentatore e il campione previene la formazione di gradienti termici locali che, in condizioni non controllate, possono influenzare il comportamento del sistema di misura. Questo miglioramento nella qualità dei dati è fondamentale per ottenere stime accurate delle proprietà meccaniche del materiale, che verranno calcolate nelle fasi successive dell’analisi.

## 2.2 Calcolo dei Parametri del Modello mediante Verosimiglianza

La caratterizzazione meccanica dei materiali mediante nanoindentazione richiede l'estrazione di parametri quantitativi dalle curve sperimentali forza-profondità. Per ottenere ciò è stata sviluppata una metodologia computazionale basata sulla massimizzazione della funzione di verosimiglianza per stimare i parametri del modello attraverso un approccio probabilistico robusto.

La scelta di un approccio probabilistico nasce dalla necessità di gestire l'incertezza intrinseca nelle misure sperimentali e di fornire stime statisticamente robuste dei parametri del modello. A differenza dei metodi tradizionali di fitting basati sui minimi quadrati, la massimizzazione della verosimiglianza offre un framework teorico più solido per la stima parametrica, permettendo inoltre di stimare simultaneamente l'incertezza delle misure attraverso il parametro  $\sigma$ .

### 2.2.1 Sviluppo del Modello Matematico

L'implementazione distingue due tipologie di comportamento meccanico durante le diverse fasi del ciclo di nanoindentazione:

#### Modello Power Law per le Fasi Dinamiche

Per le fasi di carico (loading), scarico (unloading) e scarico finale (unload\_light), il comportamento meccanico è modellato attraverso la relazione:

$$F = a \cdot (h - h_0)^m \quad (2.1)$$

dove  $F$  rappresenta la forza applicata espressa in millinewton [mN],  $h$  la profondità di penetrazione dell'indentatore misurata in nanometri [nm], mentre  $a$ ,  $h_0$  e  $m$  sono i parametri caratteristici del materiale che devono essere stimati.

Il parametro  $a$  incorpora informazioni sulle proprietà meccaniche del materiale, fungendo da fattore di scala che determina l'ampiezza della risposta meccanica. Il parametro  $h_0$  rappresenta la profondità di offset, necessaria per tenere conto del contatto iniziale tra indentatore e superficie. L'esponente  $m$  caratterizza il comportamento plastico del materiale, con valori tipici che variano a seconda della natura del materiale e della fase del ciclo di indentazione.

### Modello Costante per le Fasi di Mantenimento

Per le fasi di mantenimento (hold1 e hold2), dove la forza viene mantenuta costante nel tempo, è stato implementato un modello semplificato:

$$F = c \quad (2.2)$$

dove  $c$  rappresenta il valore costante della forza durante la fase di mantenimento.

### 2.2.2 Funzione di Verosimiglianza

La metodologia sviluppata si basa sull'assunzione che gli errori di misura seguano una distribuzione normale. Considerando che ogni segmento  $s$  del ciclo di nanoindentazione è caratterizzato da parametri specifici, la funzione di log-verosimiglianza per il segmento  $s$  è definita come:

#### Modello Power Law per i Segmenti Dinamici

Per i segmenti dinamici (loading, unloading, unload\_light), la log-verosimiglianza è:

$$\ln L_s(\theta_s) = -\frac{n_s}{2} \ln(2\pi\sigma_s^2) - \frac{1}{2\sigma_s^2} \sum_{i=1}^{n_s} (F_{s,i} - f_s(h_{s,i}, \theta_s))^2 \quad (2.3)$$

dove  $\theta_s = (a_s, h_{0,s}, m_s, \sigma_s)$  rappresenta il vettore dei parametri specifici per il segmento  $s$ ,  $n_s$  è il numero di punti sperimentali nel segmento  $s$ , e  $f_s(h_{s,i}, \theta_s) = a_s \cdot (h_{s,i} - h_{0,s})^{m_s}$ .

#### Modello Costante per i Segmenti di Mantenimento

Per i segmenti di mantenimento (hold1, hold2), la log-verosimiglianza diventa:

$$\ln L_s(\theta_s) = -\frac{n_s}{2} \ln(2\pi\sigma_s^2) - \frac{1}{2\sigma_s^2} \sum_{i=1}^{n_s} (F_{s,i} - c_s)^2 \quad (2.4)$$

dove  $\theta_s = (c_s, \sigma_s)$  e  $c_s$  rappresenta il valore costante della forza per il segmento  $s$ .

#### Ottimizzazione Indipendente per Segmento

L'ottimizzazione viene condotta separatamente per ogni segmento, massimizzando la rispettiva log-verosimiglianza:

$$\hat{\theta}_s = \arg \max_{\theta_s} \ln L_s(\theta_s) \quad (2.5)$$

Questo approccio consente di ottenere parametri ottimali specifici per ogni fase del ciclo di nanoindentazione, riflettendo le diverse caratteristiche meccaniche manifestate durante le varie fasi del processo.

### 2.2.3 Strategia di Ottimizzazione: Differential Evolution

L’ottimizzazione della funzione di verosimiglianza è implementata utilizzando l’algoritmo di Differential Evolution, una metaeuristica di ottimizzazione globale particolarmente efficace per problemi di ottimizzazione non lineare con spazi di ricerca complessi.

Il Differential Evolution appartiene alla famiglia degli algoritmi evolutivi e si basa su tre operazioni fondamentali:

- **Mutazione:** Generazione di vettori candidati attraverso combinazioni lineari di individui della popolazione
- **Crossover:** Ricombinazione dei parametri tra il vettore candidato e l’individuo corrente
- **Selezione:** Selezione dell’individuo migliore per la generazione successiva

La strategia di mutazione implementata segue il pattern DE/rand/1, dove un vettore candidato  $v_{i,g+1}$  è generato secondo:

$$v_{i,g+1} = x_{r1,g} + F \cdot (x_{r2,g} - x_{r3,g}) \quad (2.6)$$

dove  $x_{r1,g}$ ,  $x_{r2,g}$ ,  $x_{r3,g}$  sono tre individui casuali distinti della popolazione alla generazione  $g$ , e  $F$  è il fattore di mutazione.

#### Vantaggi per la Stima Parametrica

L’algoritmo di Differential Evolution presenta caratteristiche particolarmente vantaggiose per il problema della stima parametrica in nanoindentazione:

- **Robustezza ai minimi locali:** La natura stocastica dell’algoritmo consente di esplorare efficacemente lo spazio dei parametri, riducendo significativamente il rischio di convergenza a minimi locali
- **Gestione di vincoli:** L’algoritmo gestisce naturalmente i vincoli sui parametri attraverso la definizione di bounds

- **Convergenza globale:** La strategia di ricerca globale garantisce una maggiore probabilità di raggiungere il massimo globale della funzione di verosimiglianza
- **Parallelizzabilità:** La valutazione delle funzioni obiettivo per diversi individui può essere parallelizzata, migliorando l'efficienza computazionale

#### 2.2.4 Definizione dei Vincoli e Inizializzazione

Per garantire la stabilità numerica e la plausibilità fisica dei parametri stimati, vengono definiti vincoli appropriati per ciascun parametro:

##### Vincoli per il Modello Power Law

$$a \in (10^{-6}, F_{\text{range}} \times 100) \quad (2.7)$$

$$h_0 \in (0, \min(h) \times 0.9) \quad (2.8)$$

$$m \in (0.1, 5.0) \quad (2.9)$$

$$\sigma \in (10^{-6}, \sigma_F \times 2) \quad (2.10)$$

dove  $F_{\text{range}}$  è l'intervallo dei valori di forza osservati e  $\sigma_F$  è la deviazione standard empirica dei dati di forza.

L'inizializzazione dei parametri è basata su euristiche derivate dalle caratteristiche statistiche dei dati sperimentali, garantendo un punto di partenza ragionevole per l'ottimizzazione:

- Il parametro  $a$  è inizializzato considerando il rapporto tra l'intervallo di forza e una potenza caratteristica dell'intervallo di profondità
- Il parametro  $h_0$  è inizializzato come frazione del valore minimo di profondità osservato
- L'esponente  $m$  è inizializzato a un valore intermedio tipico per materiali indentati
- Il parametro  $\sigma$  è inizializzato come frazione della variabilità osservata nei dati

#### 2.2.5 Aggregazione Multi-Esperimento

La metodologia implementa un approccio innovativo che aggrega i dati da tutti gli esperimenti disponibili prima del fitting. Questa strategia presenta diversi vantaggi teorici che esponiamo a seguire.

L’aggregazione dei dati da multipli esperimenti aumenta significativamente la dimensione del campione, migliorando la potenza statistica delle stime parametriche. Con dataset aggregati contenenti migliaia di punti sperimentali, la precisione delle stime parametriche migliora secondo la legge  $\sigma_{\hat{\theta}} \propto n^{-1/2}$ , dove  $n$  è la dimensione del campione.

L’approccio aggregato riduce l’influenza di singoli esperimenti anomali o outlier, fornendo stime parametriche più robuste e rappresentative del comportamento medio del materiale.

L’aggregazione consente una stima più accurata del parametro  $\sigma$ , che cattura sia l’incertezza di misura che la variabilità intrinseca del materiale tra diversi esperimenti.

### 2.2.6 Validazione del Modello

La qualità del fitting è valutata attraverso il Root Mean Square Error (RMSE), che fornisce una misura dell’errore tipico in unità fisiche:

$$\text{RMSE}_s = \sqrt{\frac{1}{n_s} \sum_{i=1}^{n_s} (F_{s,i} - \hat{F}_{s,i})^2} \quad (2.11)$$

dove  $\hat{F}_{s,i}$  rappresenta il valore predetto dal modello per il segmento  $s$ .

#### Criteri Alternativi di Selezione del Modello

Sebbene non implementati nella presente analisi, criteri informativi alternativi come l’Akaike Information Criterion (AIC), il Deviance Information Criterion (DIC) e il Bayesian Information Criterion (BIC) potrebbero essere utilizzati per confronti più sofisticati tra modelli alternativi, bilanciando la bontà del fitting con la complessità parametrica.

## 2.3 Risultati Sperimentali

La metodologia di stima parametrica basata sulla massimizzazione della verosimiglianza è stata applicata a un esteso dataset di esperimenti di nanoindentazione. I risultati presentati in questa sezione illustrano l’efficacia dell’approccio proposto nell’estrarre parametri meccanici significativi dalle curve forza-profundità, dimostrando la robustezza statistica ottenuta attraverso l’aggregazione multi-esperimento e l’ottimizzazione globale.

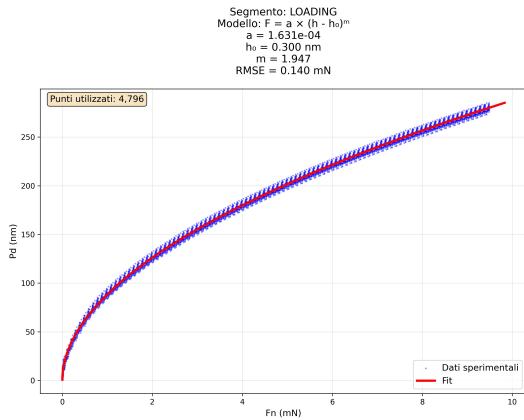


Figura 2.15. Fitting fase di Loading

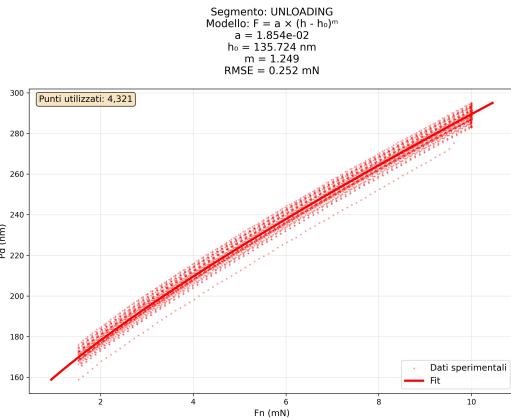


Figura 2.16. Fitting fase di unloading

Figura 2.17. Fitting significativi per dataset 1

Fase	N° Punti	RMSE	a	$h_0$	m
Loading	4796	0.140404	$1.631 \times 10^{-4}$	0.300202	1.947328
Unloading	4321	0.252283	$1.854 \times 10^{-2}$	135.724342	1.249291
Unload_Light	480	0.121776	$5.487 \times 10^{-2}$	120.388321	0.777551

Tabella 2.1. Risultati Aggregati Dataset 1 - Modelli Powerlaw

### 2.3.1 Confronto Grafico del Fitting per il Dataset 1 e 2

I grafici seguenti mostrano l'adattamento del modello ai dati sperimentali per i due dataset, rappresentativo del comportamento generale osservato in tutti i test. I grafici permettono di valutare visivamente la qualità del fitting e l'accuratezza del modello nell'interpretare i dati sperimentali.

Per entrambi i dataset proponiamo il fitting delle due fasi principali, quella di loading e unloading, presentando però in tabella i parametri per tutte e cinque le fasi, con relativi valori.

#### DataSet 1

La tabella seguente riporta i valori stimati dei parametri per il DataSet 1, con numero di punti considerati ed RMSE calcolato per il fitting.

A seguire viene mostrato anche il fit completo, diviso sempre per segmenti, del primo DataSet, che possiamo notare, come già ci aspettavamo, riprendere le figure 2.5.

Fase	N° Punti	RMSE	$c$
HOLD1	3842	0.071875	9.983396
HOLD2	28806	0.027058	1.000272

Tabella 2.2. Risultati Aggregati Dataset 1 - Modelli Costanti

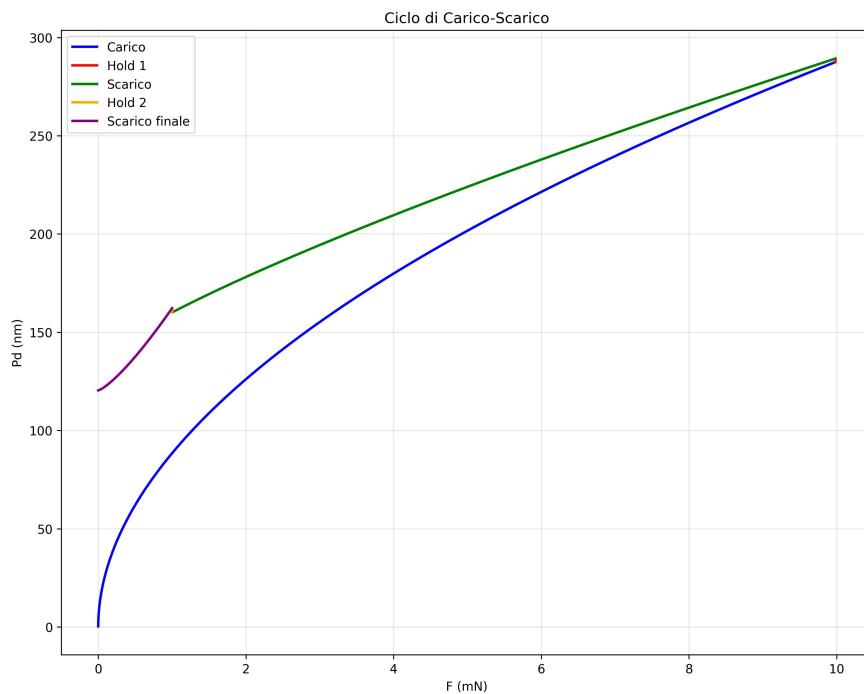


Figura 2.18. Stima del primo test per il DataSet 1

## DataSet 2

Ripetiamo lo stesso procedimento per il secondo DataSet, ovvero con l’aria condizionata accesa, in condizioni di controllo termico. Presentiamo nuovamente i fit per le due fasi più importanti di loading e unloading seguiti dalle tabelle riassuntive sui parametri.

Nuovamente come mostrato in precedenza, viene proposto il fit completo del secondo DataSet, per avere una dimostrazione pratica dei parametri appena stimati, che in questo caso riprende la figura 2.10.

I risultati ottenuti dalla procedura di stima parametrica mediante massimizzazione della verosimiglianza forniscono informazioni dettagliate e quantitative sul comportamento meccanico del materiale nelle diverse fasi del ciclo di indentazione. L’analisi comparativa tra i due dataset evidenzia l’effetto significativo delle condizioni ambientali sui parametri caratteristici e conferma l’importanza del controllo termico per la stabilità delle

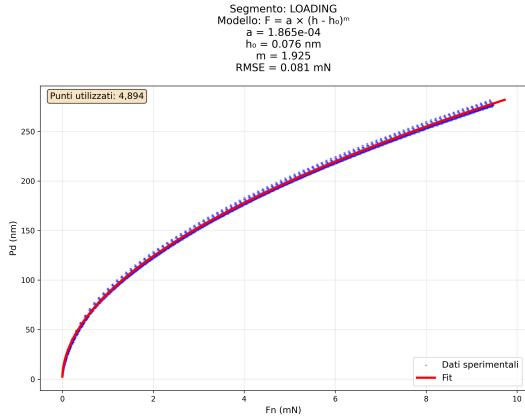


Figura 2.19. Fitting fase di Loading

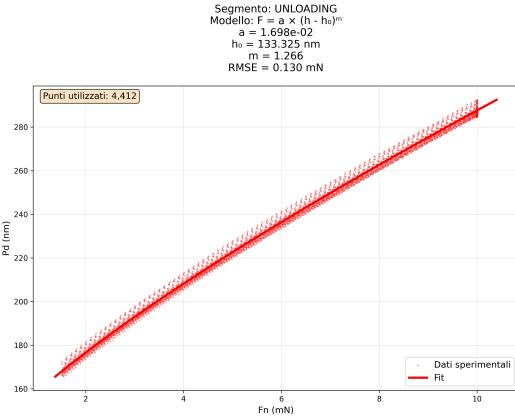


Figura 2.20. Fitting fase di unloading

Figura 2.21. Fitting significativi per DataSet 2

Fase	N° Punti	RMSE	$a$	$h_0$	$m$
Loading	4894	0.081048	$1.865 \times 10^{-4}$	0.076300	1.925362
Unloading	4412	0.130268	$1.698 \times 10^{-2}$	133.325209	1.265911
Unload_Light	490	0.104796	$5.573 \times 10^{-3}$	112.032059	1.355333

Tabella 2.3. Risultati Aggregati Dataset 2 - Modelli Powerlaw

misurazioni.

### 2.3.2 Confronto dei Parametri tra Dataset

**Fase di Loading:** Il confronto dei parametri per la fase di carico rivela differenze significative tra le due condizioni sperimentali. Il parametro  $a$  presenta un incremento nel Dataset 2 (con controllo termico) rispetto al Dataset 1, passando da  $1.631 \times 10^{-4}$  a  $1.865 \times 10^{-4}$ . Questo aumento indica una maggiore rigidezza apparente del sistema in condizioni termicamente controllate.

Il parametro  $h_0$  mostra una riduzione drastica nel Dataset 2, passando da 0.300202 nm a 0.076300 nm. Questa diminuzione significativa dell'offset di profondità conferma l'effetto stabilizzante del controllo termico, che riduce la deriva termica e migliora la precisione nella determinazione del punto di primo contatto.

L'esponente  $m$  presenta una variazione minima, mantenendosi stabile intorno a 1.94, suggerendo che il comportamento intrinseco del materiale durante la fase di carico non è significativamente influenzato dalle condizioni

Fase	N° Punti	RMSE	$c$
HOLD1	3923	0.073933	9.983785
HOLD2	29405	0.027563	1.001259

Tabella 2.4. Risultati Aggregati Dataset 2 - Modelli Costanti

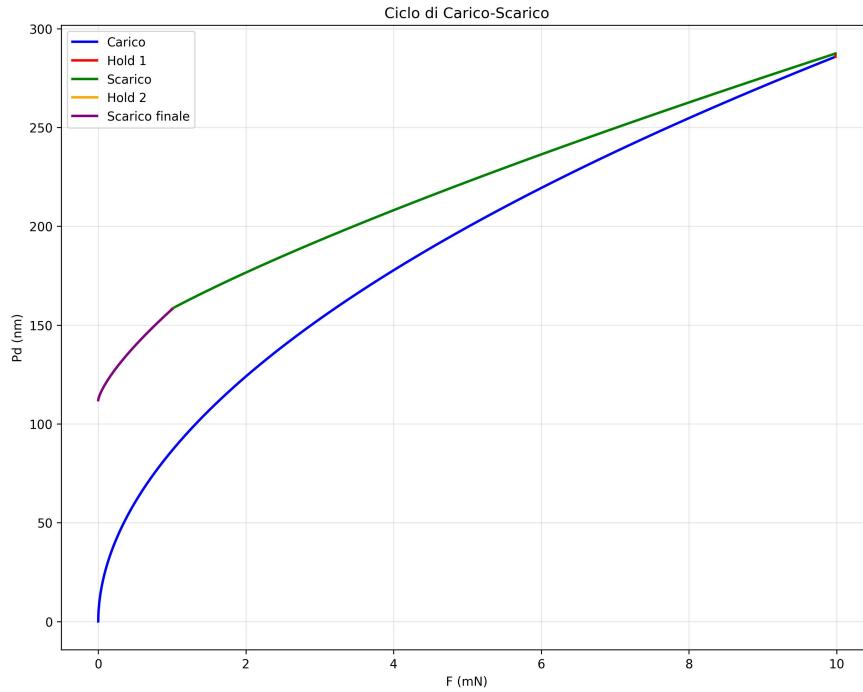


Figura 2.22. Stima del primo test per il DataSet 2

termiche.

**Fase di Unloading:** Durante la fase di scarico, il parametro  $a$  mostra una diminuzione nel Dataset 2, passando da  $1.854 \times 10^{-2}$  a  $1.698 \times 10^{-2}$ . Questo comportamento opposto rispetto alla fase di carico suggerisce che il controllo termico influenza diversamente i meccanismi di deformazione elastica e plastica.

Il parametro  $h_0$  presenta una riduzione minima, indicando una maggiore stabilità dimensionale durante il recupero elastico in condizioni termicamente controllate, mentre l'esponente  $m$  aumenta leggermente.

**Fase di Unload\_Light:** La fase di scarico finale mostra le differenze più marcate. Il parametro  $a$  diminuisce drasticamente, da  $5.487 \times 10^{-2}$  a  $5.573 \times 10^{-3}$ . Questa riduzione significativa indica una maggiore sensibilità di questa fase alle condizioni termiche.

Il parametro  $h_0$  diminuisce, mentre l'esponente  $m$  aumenta considerevolmente, passando da 0.777551 a 1.355333. Quest'ultimo risultato suggerisce un cambiamento sostanziale nel meccanismo di deformazione durante la fase finale di scarico.

### 2.3.3 Qualità del Fitting e Precisione delle Stime

L'analisi degli errori RMSE fornisce una valutazione quantitativa della qualità del fitting. Il Dataset 2 mostra una riduzione significativa dell'errore in tutte le fasi dinamiche:

- **Loading:** riduzione del 42.3% (da 0.140404 a 0.081048 mN)
- **Unloading:** riduzione del 48.4% (da 0.252283 a 0.130268 mN)
- **Unload\_Light:** riduzione del 13.9% (da 0.121776 a 0.104796 mN)

Questa riduzione sistematica degli errori nel Dataset 2 conferma che il controllo termico migliora significativamente la qualità e la riproducibilità delle misurazioni, riducendo il rumore di misura e aumentando la precisione delle stime parametriche.

### 2.3.4 Stabilità delle Fasi di Mantenimento

Le fasi di mantenimento (HOLD1 e HOLD2) mostrano parametri costanti  $c$  molto simili tra i due dataset, con variazioni inferiori allo 0.1%. Questo risultato conferma la stabilità del sistema di controllo della forza e l'efficacia del modello costante per queste fasi.

Gli errori RMSE per le fasi di mantenimento sono comparabili tra i due dataset, suggerendo che il controllo termico ha un impatto limitato sulla stabilità della forza durante le fasi statiche, mentre è particolarmente efficace durante le fasi dinamiche di carico e scarico.

### 2.3.5 Implicazioni per la Caratterizzazione Meccanica

I risultati ottenuti hanno importanti implicazioni per la caratterizzazione meccanica dei materiali:

1. **Accuratezza delle Proprietà Meccaniche:** La riduzione significativa del parametro  $h_0$  nel Dataset 2 indica una maggiore accuratezza nella determinazione del punto di primo contatto, fondamentale per il calcolo corretto delle proprietà meccaniche come durezza e modulo elastico.

2. **Riproducibilità:** La riduzione sistematica degli errori RMSE nel Dataset 2 suggerisce una maggiore riproducibilità delle misurazioni in condizioni termicamente controllate, aspetto cruciale per applicazioni metrologiche.
3. **Sensibilità alle Condizioni Ambientali:** Le differenze sostanziali nei parametri tra i due dataset evidenziano l'importanza critica del controllo delle condizioni ambientali per ottenere risultati affidabili e confrontabili.

# Capitolo 3

## Analisi Bayesiana con Stan per la Nanoindentazione

### 3.1 Introduzione a Stan e all’Analisi Bayesiana

Stan è un linguaggio di programmazione probabilistico sviluppato per l’inferenza statistica Bayesiana e l’ottimizzazione. A differenza dei metodi di massima verosimiglianza utilizzati nel precedente capitolo, l’approccio Bayesiano permette di incorporare conoscenze a priori sui parametri e di quantificare l’incertezza attraverso distribuzioni di probabilità a posteriori.

#### 3.1.1 Il Framework Bayesiano

Il framework Bayesiano si basa sul teorema di Bayes:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \quad (3.1)$$

dove:

- $P(\theta|D)$  è la distribuzione a posteriori dei parametri  $\theta$  dati i dati  $D$
- $P(D|\theta)$  è la verosimiglianza
- $P(\theta)$  è la distribuzione a priori
- $P(D)$  è l’evidenza marginale

### 3.1.2 Algoritmo Hamiltonian Monte Carlo (HMC)

Stan utilizza l'algoritmo Hamiltonian Monte Carlo (HMC) con algoritmo No-U-Turn Sampler (NUTS) per campionare dalla distribuzione a posteriori. L'HMC sfrutta informazioni del gradiente per esplorare efficacemente lo spazio dei parametri, particolarmente vantaggioso per modelli ad alta dimensionalità.

L'algoritmo HMC introduce una variabile momento ausiliaria  $p$  e definisce un'energia potenziale  $U(\theta) = -\log P(\theta|D)$  e un'energia cinetica  $K(p) = \frac{1}{2}p^T M^{-1}p$ . Il sistema Hamiltoniano risultante è:

$$\frac{d\theta}{dt} = M^{-1}p \quad (3.2)$$

$$\frac{dp}{dt} = -\frac{\partial U}{\partial \theta} \quad (3.3)$$

Il No-U-Turn Sampler (NUTS) determina automaticamente la lunghezza ottimale delle traiettorie, eliminando la necessità di tuning manuale del numero di passi di integrazione.

## 3.2 Preparazione e Caricamento dei Dati

I dati di nanoindentazione sono organizzati in file di testo contenenti misurazioni temporali di

- **Tempo (s)**: timestamp delle misurazioni
- **Profondità (Pd)**: penetrazione dell'indentatore in nanometri
- **Forza normale (Fn)**: forza applicata in millinewton
- **ID Segmento**: identificatore del segmento del ciclo di indentazione (1-5)

Il processo di caricamento implementa diverse fasi di controllo qualità attraverso la funzione `load_file()`.

I criteri di filtraggio applicati garantiscono:

- Eliminazione di valori negativi non fisici per la profondità ( $Pd > 0$ )
- Rimozione di forze eccessivamente negative ( $Fn > -0.001$ )
- Validazione dei segmenti di indentazione ( $SegmentID \in \{1,2,3,4,5\}$ )
- Gestione dei valori mancanti o corrotti

A seguire viene riproposto il modello che mantiene la stessa struttura fisica del capitolo precedente, con funzioni specifiche per ogni segmento, garantendo però la continuità tra i segmenti attraverso il calcolo dinamico di  $c_1$ ,  $c_2$ ,  $h_2$  e  $h_3$ :

$$\mu_i = \begin{cases} a_1 \cdot (h_i - h_1)^{m_1} & \text{se } segment_i = 1 \text{ (loading)} \\ & \text{dove } h_1 = 0 \\ a_1 \cdot (h_k - h_1)^{m_1} = c_1 & \text{se } segment_i = 2 \text{ (hold1)} \\ a_2 \cdot (h_i - h_2)^{m_2} & \text{se } segment_i = 3 \text{ (unloading)} \\ & \text{dove } h_2 = h_i - \left(\frac{a_1 \cdot (h_k - h_1)^{m_1}}{a_2}\right)^{1/m_2} \\ & \text{se } segment_{i-1} = 2 \\ a_2 \cdot (h_k - h_2)^{m_2} = c_2 & \text{se } segment_i = 4 \text{ (hold2)} \\ a_3 \cdot (h_i - h_3)^{m_3} & \text{se } segment_i = 5 \text{ (unload\_light)} \\ & \text{dove } h_3 = h_i - \left(\frac{a_2 \cdot (h_k - h_2)^{m_2}}{a_3}\right)^{1/m_3} \\ & \text{se } segment_{i-1} = 4 \end{cases} \quad (3.4)$$

dove  $k$  è l'indice dell'ultimo punto calcolato nel segmento precedente. In particolare, si imposta  $h_1 = 0$  poiché al tempo  $t = 0$  si ha  $h = 0$  e  $F = 0$ . Considerando che  $a \neq 0$ , ciò implica direttamente la condizione iniziale stabilità. Successivamente, le costanti  $c_1$  e  $c_2$  vengono determinate utilizzando i valori dell'ultimo punto del segmento precedente alla fase di hold, garantendo così la continuità nei punti di inizio delle fasi di mantenimento. Infine, i parametri  $h_2$  e  $h_3$  si ottengono applicando la condizione di uguaglianza tra le equazioni delle fasi di *loading – unloading* e *unloading – unload\_light*. Questo approccio assicura che l'ultimo valore della fase precedente, mantenuto costante durante la fase di hold, coincida con il primo valore della nuova fase, garantendo nuovamente la continuità della funzione.

La verosimiglianza del modello è definita come:

$$F \sim \text{Normal}(\mu, \sigma) \quad (3.5)$$

### 3.3 Implementazione in Stan

#### 3.3.1 Struttura del Modello Stan

Il modello Stan è strutturato in quattro blocchi principali:

Listing 3.1. Blocco data e parameters del modello Stan

```

data {
  int<lower=0> N;
  vector[N] h;
  vector[N] F;
  int<lower=1,upper=5> segment[N];
}

parameters {
  vector<lower=1e-8, upper=10>[3] a;
  vector<lower=0.1, upper=5>[3] m;
  real<lower=0.001, upper=50> sigma;
}

```

Listing 3.2. Blocco transformed parameters (estratto principale)

```

transformed parameters {
  vector[N] mu;
  real h1 = 0;
  real h2;
  real h3;
  real h_last;
  vector[2] c;

  for (i in 1:N) {
    if (segment[i] == 1)
    { // loading
      mu[i] = a[1] * pow(fmax(h[i] - h1, 1e-9), m[1]);
      h_last=h[i];
    }
    else if (segment[i] == 2)
    { // hold1
      mu[i] = a[1] * pow(fmax(h_last - h1, 1e-9), m[1]);
      c[1] = mu[i];
    }
    else if (segment[i] == 3)
    { // unloading
      if (segment[i-1] == 2)
      {
        h2 = h[i] - pow((a[1]*pow(fmax(h_last - h1, 1e-9), m[1]))/a[2], 1.0 / m[2]);
        mu[i] = a[2] * pow(fmax(h[i] - h2, 1e-9), m[2]);
      }
    }
  }
}

```

Listing 3.3. Blocchi model e generated quantities

```

model {
  // Likelihood
  F ~ normal(mu, sigma);
}
```

```
generated_quantities {  
    vector[N] F_rep;  
    vector[N] log_lik;  
    vector[3] h0;  
    vector[2] c_output;
```

h0[1] = h1;

```
h0[2] = h2;
h0[3] = h3;

c_output[1] = c[1];
c_output[2] = c[2];

for (i in 1:N) {
  F_rep[i] = normal_rng(mu[i], sigma);
  log_lik[i] = normal_lpdf(F[i] | mu[i], sigma);
}
}
```

### 3.3.2 Strategie di Inizializzazione e Parametri

L'inizializzazione dei parametri è critica per la convergenza dell'algoritmo NUTS. È stata sviluppata una funzione di inizializzazione che:

- Utilizza seed diversi per ogni catena
- Inizializza in regioni ad alta probabilità
- Rispetta rigorosamente i bounds definiti

Listing 3.4. Funzione di inizializzazione

```
init <- function(chain_id = 1) {
  set.seed(12345 + chain_id * 100)

  list(
    a = c(
      runif(1, 1e-6, 1e-2),      # loading
      runif(1, 1e-4, 1e-1),      # unloading
      runif(1, 1e-4, 1e-1)       # unload_light
    ),
    m = c(
      runif(1, 1.5, 2.5),      # loading
      runif(1, 0.8, 1.5),      # unloading
      runif(1, 0.5, 1.2)       # unload_light
    ),
    sigma = runif(1, 0.1, 2)
  )
}
```

La configurazione MCMC adottata bilancia accuratezza e efficienza computazionale:

- **Catene:** 4 catene indipendenti per diagnostica di convergenza
- **Iterazioni:** 4000 totali (2000 di campionamento)
- **Warmup :** 2000 iterazioni
- **Thinning:** 2
- **Cores:** 4 (parallelizzazione)

Listing 3.5. Configurazione del fit Stan

```
fit <- stan(
  file = model_file ,
  data = stan_data ,
  init = init_list ,
  chains = 4 ,
  iter = 4000 ,
  warmup = 2000 ,
  thin = 2 ,
  cores = 4 ,
  verbose = FALSE
)
```

## 3.4 Diagnostiche di Convergenza

### 3.4.1 R-hat Statistic ( $\hat{R}$ )

L' $\hat{R}$  (R-hat) quantifica la convergenza confrontando la varianza tra catene con quella all'interno delle catene:

$$\hat{R} = \sqrt{\frac{\hat{V}}{W}} \quad (3.6)$$

dove  $\hat{V}$  è la stima pooled della varianza e  $W$  è la varianza media intra-catena.

**Criteri di interpretazione:**

- $\hat{R} < 1.01$ : Convergenza eccellente
- $1.01 \leq \hat{R} < 1.1$ : Convergenza accettabile

### 3.4.2 Effective Sample Size (ESS)

L'ESS quantifica il numero di campioni indipendenti equivalenti, considerando l'autocorrelazione:

$$\text{ESS} = \frac{MN}{1 + 2 \sum_{t=1}^T \rho_t} \quad (3.7)$$

dove  $M$  è il numero di catene,  $N$  il numero di iterazioni post-warmup, e  $\rho_t$  è l'autocorrelazione al lag  $t$ .

**Soglie raccomandate:**

- **ESS bulk** > 400: Stima affidabile della media a posteriori
- **ESS tail** > 400: Stima affidabile delle code della distribuzione

Tabella 3.1. ESS e Rhat per parametri chiave - Dataset 1

Parametro	ESS bulk	ESS tail	Rhat
a[1]	3011.4870	2889.8237	1.0015
m[1]	3003.6543	2868.0459	1.0018
a[2]	3563.7823	3380.8394	1.0009
m[2]	3562.6805	3322.1902	1.0009
a[3]	9.3846	15.7683	1.2475
m[3]	9.3758	14.8960	1.4402
sigma	3049.1717	3113.7022	0.9998
$h_0[1]$	NA	NA	NA
$h_0[2]$	3568.5798	3440.8380	1.0009
$h_0[3]$	9.2616	16.9649	1.4163
c[1]	3585.1613	3340.4434	0.9998
c[2]	3838.8946	3691.6698	1.0002

## 3.5 Risultati del Modello

In questa sezione vengono presentati i risultati del modello STAN descritto precedentemente. In particolare, per ciascun parametro, riportiamo inizialmente le diagnostiche di convergenza, seguite dalla stima puntuale dei parametri con i relativi intervalli di credibilità. Infine, viene riportato il valore dell'RMSE del fit generato utilizzando i parametri stimati, al fine di consentire un confronto diretto con la stima ottenuta nel capitolo precedente mediante Python.

La notazione utilizzata nelle tabelle seguenti è la seguente:

- Per i parametri  $a$ ,  $m$  e  $h_0$ :
  - 1 indica la fase di *loading*,
  - 2 la fase di *unloading*,
  - 3 la fase di *unload\_light*.
- Per il parametro  $c$ :
  - 1 si riferisce alla prima fase di *hold*,
  - 2 alla seconda fase di *hold*.

### 3.5.1 Diagnostiche per Parametro - Dataset 1

Come si può osservare nella Tabella 3.1, l'Effective Sample Size (ESS), sia in riferimento alla media a posteriori che alla varianza, risulta elevato per

Tabella 3.2. ESS e Rhat per parametri chiave - Dataset 2

Parametro	ESS bulk	ESS tail	Rhat
a[1]	1095.0142	1324.3937	1.0001
m[1]	1091.2205	1393.9030	1.0002
a[2]	3365.2230	3487.9871	0.9998
m[2]	3333.4500	3487.9871	0.9999
a[3]	6.8074	14.0844	2.0891
m[3]	6.9256	13.1961	1.3553
sigma	3084.6593	3589.9986	1.0004
$h_0[1]$	NA	NA	NA
$h_0[2]$	3470.8104	3520.2913	0.9999
$h_0[3]$	7.1353	13.4793	1.3755
c[1]	2487.8152	3233.9011	1.0003
c[2]	3841.6321	3661.2810	1.0001

la maggior parte dei parametri. Fanno eccezione i parametri  $a$ ,  $m$  e  $h_0$  relativi alla fase di *unload\_light*, che presentano valori di ESS più bassi e, contemporaneamente, un valore di  $\hat{R}$  maggiore rispetto agli altri.

Per quanto riguarda il parametro  $h_0$  relativo alla fase di *loading*, esso presenta valori NA in quanto è stato deterministicamente fissato a 0, come indicato nel Modello 3.4. Di conseguenza, per tale parametro non sono disponibili né l'ESS né il valore di  $\hat{R}$ , in quanto non è stato oggetto di campionamento durante la stima bayesiana.

### 3.5.2 Diagnostiche per Parametro - Dataset 2

Confrontando con la Tabella 3.1, si osservano valori di ESS e  $\hat{R}$  molto simili a quelli precedentemente riportati, indicando che le catene hanno raggiunto una buona convergenza e che la stima bayesiana risulta complessivamente affidabile.

Anche in questo caso, tuttavia, si riscontra una situazione critica per i parametri associati alla fase di *unload\_light*, che mostrano valori di  $\hat{R}$  significativamente più alti e, di contro, valori di ESS più bassi. Questo comportamento persiste nonostante l'introduzione del controllo termico, confermando la maggiore difficoltà nella stima di tali parametri, probabilmente dovuta alla maggiore variabilità della fase stessa.

Tabella 3.3. Stime a posteriori Bayesiane - Dataset 1

Parametro	Mean	SE_Mean	SD	2.5%	97.5%
a[1]	0.0001	2.71E-08	1.50E-06	9.77E-05	0.0001
m[1]	2.0420	4.83E-05	0.0027	2.0368	2.0473
a[2]	0.0247	5.57E-06	0.0003	0.0241	0.0254
m[2]	1.2138	4.30E-05	0.0026	1.2088	1.2189
a[3]	0.0057	0.0059	0.0092	1.91E-06	0.0297
m[3]	2.0921	0.3711	0.6197	1.1730	3.3539
sigma	0.2222	1.36E-05	0.0008	0.2207	0.2237
$h_0[1]$	0	NA	0	0	0
$h_0[2]$	149.3808	0.0018	0.1095	149.1668	149.5953
$h_0[3]$	123.9432	5.1952	8.7893	106.0518	137.0321
c[1]	9.6912	4.68E-05	0.0028	9.6858	9.6967
c[2]	0.9186	2.27E-05	0.0014	0.9158	0.9215

### 3.5.3 Stime a Posteriori dei Parametri - Dataset 1

In questa tabella sono riportate le stime di tutti i parametri del modello, ricordando che il parametro  $h_0$  relativo alla fase di *loading* è fissato a zero. I risultati ottenuti sono coerenti con quanto osservato nella tabella precedente, sempre riferita al Dataset 1: anche in questo caso, si riscontrano valori molto bassi sia per l'errore standard (SE) che per la deviazione standard (SD), accompagnati da intervalli di credibilità al 95% particolarmente stretti.

Fanno eccezione i tre parametri relativi alla fase di *unload\_light*, i quali presentano valori più elevati di SE, SD e intervallo di credibilità, confermando una maggiore variabilità e una stima meno precisa.

### 3.5.4 Stime a Posteriori dei Parametri - Dataset 2

Confrontando con la Tabella 3.3 e in accordo con quanto evidenziato nella Tabella 3.2, i risultati ottenuti per il secondo dataset sono complessivamente simili a quelli del primo. Per la maggior parte dei parametri si osservano deviazioni standard (SD) ed errori standard (SE) contenuti, accompagnati da intervalli di credibilità al 95% ristretti, a indicare una buona precisione delle stime.

Fanno eccezione, ancora una volta, i parametri  $a$ ,  $m$  e  $h_0$  relativi all'ultima fase, coerentemente con quanto già osservato nel primo dataset.

Tabella 3.4. Stime a posteriori Bayesiane - Dataset 2

Parametro	Mean	SE_Mean	SD	2.5%	97.5%
a[1]	9.65E-05	2.89E-08	9.56E-07	9.47E-05	9.84E-05
m[1]	2.0514	5.37E-05	0.0018	2.0479	2.0549
a[2]	0.0511	7.10E-06	0.0004	0.0503	0.0519
m[2]	1.0746	2.72E-05	0.0016	1.0715	1.0776
a[3]	0.0323	0.0223	0.0316	9.78E-07	0.0785
m[3]	1.6057	0.5280	0.7955	0.8818	3.4398
sigma	0.1463	9.35E-06	0.0005	0.1454	0.1473
$h_0[1]$	0	NA	0	0	0
$h_0[2]$	152.8310	0.0010	0.0597	152.7165	152.9464
$h_0[3]$	129.5004	7.9760	12.0681	101.0891	140.7665
c[1]	9.8336	3.78E-05	0.0019	9.8298	9.8372
c[2]	0.9963	1.38E-05	0.0009	0.9946	0.9980

### 3.5.5 Confronto degli Intervalli di Credibilità tra Dataset

Per valutare la consistenza dei parametri stimati tra i due dataset, è fondamentale verificare se gli intervalli di credibilità al 95% si sovrappongono. La sovrapposizione degli intervalli fornisce evidenza empirica sulla equivalenza statistica dei parametri, mentre la loro separazione indica differenze significative tra le condizioni sperimentali.

L'analisi degli intervalli di credibilità rivela un quadro eterogeneo di consistenza tra i due dataset:

- **Fase di Loading:** Il parametro  $a[1]$  mostra piena sovrapposizione, indicando comportamento elastico consistente. Tuttavia, il parametro  $m[1]$  presenta una sovrapposizione solo marginale, con l'estremo superiore del Dataset 1 (2.0473) che coincide quasi esattamente con l'estremo inferiore del Dataset 2 (2.0479), suggerendo una possibile differenza nelle proprietà del materiale durante la fase di carico.
- **Fase di Unloading:** Si osserva una separazione completa degli intervalli per tutti e tre i parametri ( $a[2]$ ,  $m[2]$ ,  $h_0[2]$ ). Questa evidenza statistica indica che il controllo termico ha influenzato significativamente il comportamento meccanico durante la fase di scarico. In particolare, il valore di  $a[2]$  risulta circa doppio nel Dataset 2, mentre  $m[2]$  è considerevolmente inferiore, suggerendo un diverso meccanismo di rilascio elastico.

Tabella 3.5. Confronto degli intervalli di credibilità al 95% tra Dataset 1 e Dataset 2

Parametro	Dataset 1 (95% CI)	Dataset 2 (95% CI)	Sovrapposizione
<b>Fase Loading</b>			
a[1]	[9.77E-05, 1.00E-04]	[9.47E-05, 9.84E-05]	Sì
m[1]	[2.0368, 2.0473]	[2.0479, 2.0549]	No
<b>Fase Unloading</b>			
a[2]	[0.0241, 0.0254]	[0.0503, 0.0519]	No
m[2]	[1.2088, 1.2189]	[1.0715, 1.0776]	No
$h_0[2]$	[149.1668, 149.5953]	[152.7165, 152.9464]	No
<b>Fase Unload_Light</b>			
a[3]	[1.91E-06, 0.0297]	[9.78E-07, 0.0785]	Sì
m[3]	[1.1730, 3.3539]	[0.8818, 3.4398]	Sì
$h_0[3]$	[106.0518, 137.0321]	[101.0891, 140.7665]	Sì
<b>Fasi Hold</b>			
c[1]	[9.6858, 9.6967]	[9.8298, 9.8372]	No
c[2]	[0.9158, 0.9215]	[0.9946, 0.9980]	No

- **Fase di Unload\_Light:** Nonostante l'elevata incertezza evidenziata dai valori di  $\hat{R}$  e ESS, tutti i parametri mostrano sovrapposizione degli intervalli. Questo risultato è frutto però di intervalli di credibilità ben più ampi nel caso dei 2 parametri della rispettiva fase, ma indica una uguale difficoltà nel stimarli per quanto riguarda entrambi i dataset.
- **Fasi di Hold:** Entrambi i parametri  $c[1]$  e  $c[2]$  presentano intervalli completamente separati, con valori sistematicamente più elevati nel Dataset 2. Questo comportamento suggerisce che il controllo termico ha modificato le proprietà viscoelastiche del materiale durante le fasi di mantenimento del carico.

La sintesi di questa analisi evidenzia che il controllo termico implementato nel Dataset 2 ha prodotto effetti differenziali sulle varie fasi del processo di nanoindentazione. Le modifiche più pronunciate si osservano nelle fasi di *unloading* e *hold*, mentre le fasi di *loading* e *unload\_light* mantengono una maggiore consistenza, seppur con alcune eccezioni. Questo pattern suggerisce che il controllo termico influenzi principalmente i meccanismi di rilassamento e recupero elastico, piuttosto che le proprietà fondamentali di deformazione plastica del materiale.

Tabella 3.6. RMSE per fase - Dataset 1

Fase	RMSE	N. Punti
Loading	0.2611	4796
Hold1	0.3009	3842
Unloading	0.4233	4321
Hold2	0.0860	28806
Unload_Light	7.5320	480
<b>Totale</b>	<b>0.8236</b>	<b>42245</b>

### 3.5.6 Criteri di Validazione tramite RMSE

La performance del modello Bayesiano viene valutata confrontando l’RMSE ottenuto con i risultati dell’analisi frequentista precedentemente condotta in Python. Come mostrato nelle tabelle del capitolo precedente, l’approccio frequentista ha prodotto RMSE di riferimento per entrambi i dataset.

L’RMSE viene calcolato sia globalmente per l’intero dataset che per ciascuna fase del processo di nanoindentazione, permettendo di identificare eventuali problematiche specifiche in determinate fasi del ciclo.

Il confronto permette di valutare se l’approccio Bayesiano mantiene o migliora la qualità del fit, considerando che utilizza distribuzioni a priori e un diverso algoritmo di stima.

### 3.5.7 Metriche di Performance - Dataset 1

Nell’ultima tabella riportiamo, per il primo dataset, le metriche di performance del modello per ciascun segmento: RMSE e numero di punti utilizzati. In fondo alla tabella sono inoltre presentati i valori complessivi calcolati sull’intero dataset.

Poiché il modello è stato costruito in modo da garantire la continuità tra i diversi segmenti, è ragionevole aspettarsi prestazioni leggermente inferiori a livello di singolo tratto rispetto ai modelli separati presentati nelle Tabelle 2.1 e 2.2. Tuttavia, questa scelta consente un significativo miglioramento della coerenza globale del modello, come evidenziato dai valori complessivi più stabili.

Coerentemente con le analisi precedenti, la fase di *unload\_light* presenta le peggiori prestazioni, con un RMSE sensibilmente più elevato.

Oltre ai semplici valori di RMSE per le varie fasi, proponiamo anche il grafico con i valori di forza ottenuti al variare delle profondità in maniera

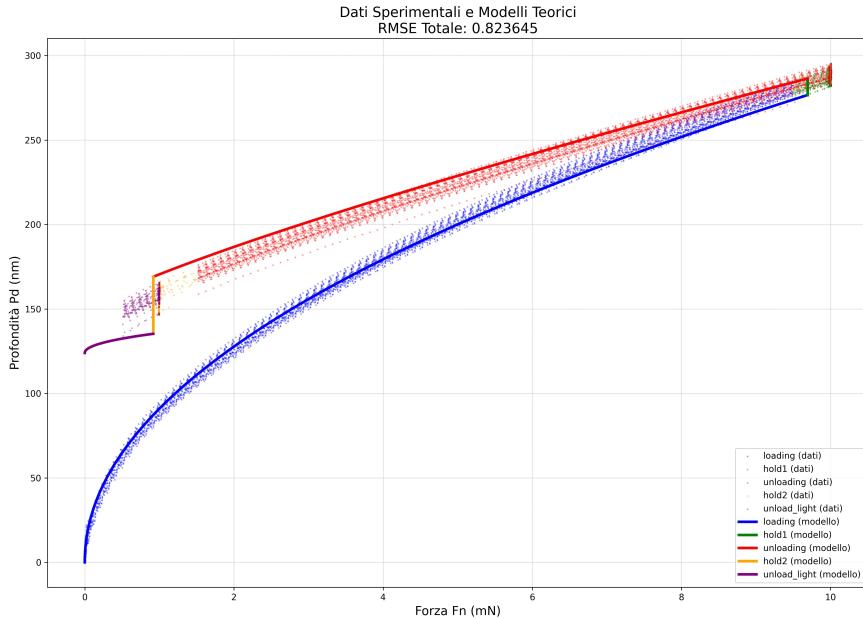


Figura 3.1. Fit con parametri modello normale per Dataset 1

Tabella 3.7. RMSE per fase - Dataset 2

Fase	RMSE	N. Punti
Loading	0.2131	4894
Hold1	0.1674	3923
Unloading	0.2421	4412
Hold2	0.0280	29405
Unload_Light	4.7499	490
<b>Totale</b>	<b>0.5200</b>	<b>43124</b>

da avere un rappresentazione chiara ed efficace dei parametri che siamo andati a stimare.

Dalla figura 3.1, possiamo notare come il modello normale, per il primo dataset, presenti degli ottimi fit per tutte le fasi, ad esclusione dell'ultima, ovvero *unload\_light*, come però ci aspettavamo già precedentemente, avendo analizzato i risultati di convergenza e i relativi RMSE.

### 3.5.8 Metriche di Performance - Dataset 2

In questa tabella emergono con maggiore evidenza le differenze tra i due dataset. Rispetto alla Tabella 3.6, si osserva infatti un RMSE inferiore in tutte le fasi, con una conseguente riduzione anche dell'RMSE totale.

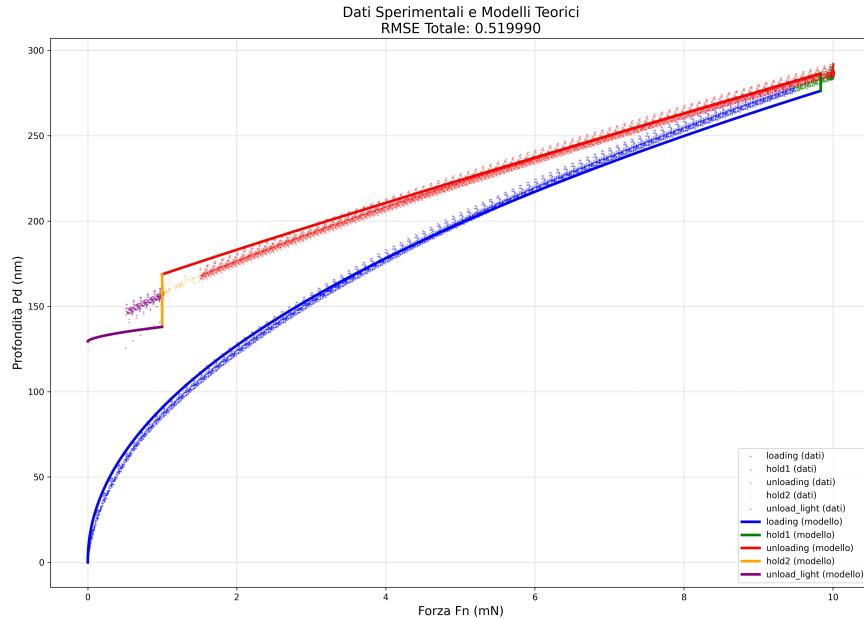


Figura 3.2. Fit con parametri modello normale per Dataset 2

Per le fasi di *loading* e *unloading*, le differenze risultano contenute, in quanto i fit erano già di buona qualità nel primo dataset. Tuttavia, per le fasi di *hold* e, in particolare, *unload\_light*, si nota un miglioramento sostanziale.

Tale miglioramento può essere attribuito sia alla presenza di un test aggiuntivo, che ha incrementato il numero di punti disponibili per la stima, sia all’impiego del controllo termico tramite aria condizionata durante l’acquisizione dei dati del secondo dataset, che ha presumibilmente ridotto la variabilità delle misure.

Come fatto per il Dataset 1, proponiamo nuovamente il confronto tra i dati esperimentali e il nostro modello normale. I risultati tra i due dataset sono molto simili in quanto a fit, a meno di notare un leggero miglioramento nella fase finale, presentando comunque un ottimo risultato nel complesso.

### 3.6 Modelli Alternativi: Log-Normale e Gamma

Il modello di base presentato nelle sezioni precedenti assume che la variabile di risposta  $F$  (forza normale) segua una distribuzione normale:

$$F \sim \text{Normal}(\mu, \sigma) \quad (3.8)$$

Tuttavia, questa assunzione potrebbe non essere ottimale per dati di nanoindentazione per diverse ragioni:

- **Positività:** Le forze fisiche sono intrinsecamente positive, mentre la distribuzione normale può assumere valori negativi
- **Simmetria:** La normale è simmetrica rispetto alla media, ma i dati sperimentali potrebbero presentare asimmetrie
- **Eteroschedasticità:** La varianza degli errori potrebbe dipendere dalla magnitudine della forza misurata
- **Distribuzione degli errori:** Gli errori di misura in sistemi fisici spesso seguono distribuzioni non gaussiane

Per affrontare queste limitazioni, sono state esplorate due distribuzioni alternative che mantengono la stessa struttura del modello fisico ma modificano l'assunzione distributiva:

1. **Distribuzione Log-Normale:**  $\log(F) \sim \text{Normal}(\log(\mu), \sigma)$
2. **Distribuzione Gamma:**  $F \sim \text{Gamma}(\text{alpha}, \text{alpha}/\mu)$

#### 3.6.1 Modello Log-Normale

Nel modello log-normale, si assume che il logaritmo della forza segua una distribuzione normale:

$$\log(F_i) \sim \text{Normal}(\log(\mu_i), \sigma_{\log}) \quad (3.9)$$

Questo implica che:

$$F_i \sim \text{LogNormal}(\log(\mu_i), \sigma_{\log}) \quad (3.10)$$

Le modifiche implementative rispetto al modello base sono minimali:

- **Blocco model:** Sostituzione della likelihood normale con log-normale
- **Generated quantities:** Utilizzo di `lognormal_rng()` e `lognormal_lpdf()`
- **Parametri:** Nessuna modifica ai parametri del modello

### Vantaggi del Modello Log-Normale

- **Positività garantita:**  $F > 0$  per costruzione
- **Moltiplicatività degli errori:** Appropriato quando gli errori sono proporzionali alla magnitudine del segnale
- **Asimmetria:** Può catturare distribuzioni con coda destra più pesante
- **Semplicità:** Richiede modifiche minime al codice esistente

### 3.6.2 Modello Gamma

Nel modello Gamma, la forza segue direttamente una distribuzione Gamma:

$$F_i \sim \text{Gamma}(\text{alpha}, \text{rate}_i) \quad (3.11)$$

con parametrizzazione tale che  $\mathbb{E}[F_i] = \mu_i$ :

$$\text{rate}_i = \frac{\text{alpha}}{\mu_i} \quad (3.12)$$

Le modifiche implementative includono:

- **Parametri aggiuntivi:** Aggiunta del parametro `alpha`
- **Blocco model:** Sostituzione con likelihood Gamma
- **Generated quantities:** Utilizzo di `gamma_rng()` e `gamma_lpdf()`
- **Inizializzazioni:** Inclusione di valori iniziali per `alpha`

### Vantaggi del Modello Gamma

- **Flessibilità:** Può assumere forme diverse (esponenziale quando `alpha=1`, simil-normale per `alpha` elevati)
- **Positività:**  $F > 0$  per costruzione
- **Controllo della varianza:** Il parametro `alpha` permette di modellare esplicitamente la variabilità
- **Interpretabilità fisica:** Spesso utilizzata per modellare tempi di attesa e intensità di processi fisici

Tabella 3.8. ESS e Rhat per parametri chiave - Modello Log-Normale, Dataset 1

Parametro	ESS bulk	ESS tail	Rhat
a[1]	1884.0080	2253.6510	1.0001
m[1]	1879.1490	2188.8320	1.0005
a[2]	1731.6860	2118.7180	1.0097
m[2]	1737.3010	2075.6990	1.0096
a[3]	1.3285	14.1057	2.1016
m[3]	1.3179	14.2452	2.1017
sigma	40.6792	2310.9660	1.0285
$h_0[1]$	NA	NA	NA
$h_0[2]$	1745.9830	1964.5560	1.0095
$h_0[3]$	1.3189	14.2875	2.0993
c[1]	3178.4830	2466.2900	0.9998
c[2]	2906.3480	2586.5350	1.0007

Tabella 3.9. Stime a posteriori Bayesiane - Modello Log-Normale, Dataset 1

Parametro	Mean	SE_Mean	SD	2.5%	97.5%
a[1]	0.0002	1.41E-07	4.27E-06	0.0002	0.0002
m[1]	1.9713	0.0001	0.0042	1.9674	1.9751
a[2]	0.0001	3.40E-06	9.17E-05	0.0001	0.0001
m[2]	2.1691	0.0009	0.0235	2.1457	2.1934
a[3]	0.0050	0.0042	0.0065	3.24E-08	0.0151
m[3]	2.3024	0.6331	0.9496	1.2935	4.1003
sigma	0.1284	3.31E-05	0.0006	0.1274	0.1293
$h_0[1]$	0	NA	0	0	0
$h_0[2]$	107.1531	0.0414	1.0359	106.0220	108.2175
$h_0[3]$	115.6540	9.5746	14.3897	88.2583	131.2259
c[1]	9.8889	0.0007	0.0207	9.8633	9.9163
c[2]	0.9851	2.77E-05	0.0009	0.9835	0.9866

## 3.7 Risultati dei Modelli Alternativi

### 3.7.1 Modello Log-Normale - Dataset 1

Confrontando i risultati ottenuti nelle Tabelle 3.8, 3.9 e 3.10 con quelli delle corrispondenti tabelle del modello originale (Tabelle 3.1, 3.3 e 3.6), non si osserva un miglioramento significativo nella qualità delle stime dei

Tabella 3.10. RMSE per fase - Modello Log-Normale, Dataset 1

Fase	RMSE	N. Punti
Loading	0.4034	4796
Hold1	0.1187	3842
Unloading	0.7759	4321
Hold2	0.0310	28806
Unload_Light	25.0510	480
<b>Totale</b>	<b>2.7979</b>	<b>42245</b>

parametri.

Al contrario, per la maggior parte delle fasi i risultati risultano leggermente peggiorati, con un peggioramento particolarmente marcato per i parametri della fase di *unload\_light*, che mostrano un ESS ulteriormente ridotto, valori di SE e SD più elevati e un picco anomalo nell'RMSE.

Anche l'RMSE totale risulta sensibilmente superiore rispetto a quello ottenuto con il modello originale, suggerendo che, almeno per il Dataset 1, questo approccio alternativo non risulti ottimale.

Ugualmente a quanto fatto per il modello Normale, proponiamo i grafici per entrambi i Dataset della funzione ottenuta con i parametri dalle stime a posteriori e i dati iniziali, per avere un riscontro grafico del lavoro appena concluso.

Come si evince molto chiaramente dal grafico il modello Log-Normale, questo modello non ha portato risultati ottimali, il che ci era già stato indicato dai valori di RMSE trovati precedente. In particolare, possiamo notare come il picco anomale di RMSE citato in analisi alle tabelle, si riscontri in uno discostamento totale della nostra predizione rispetto ai dati originali, confermando le ipotesi precedenti. Inoltre se confrontato con 3.1, notiamo che la differenza è netta, con il modello Normale che segue molto più coerentemente l'andamento dei dati a nostra disposizione.

### 3.7.2 Modello Log-Normale - Dataset 2

Come discusso precedentemente per il modello Normale, il Dataset 2 presenta una stima dei parametri decisamente migliore rispetto a quella ottenuta con il primo dataset, a parità di modello. Si osservano miglioramenti rilevanti in quasi tutte le fasi, ad eccezione delle fasi di *hold 1* e *hold 2*, che non mostrano variazioni significative, essendo già relativamente semplici da modellare.

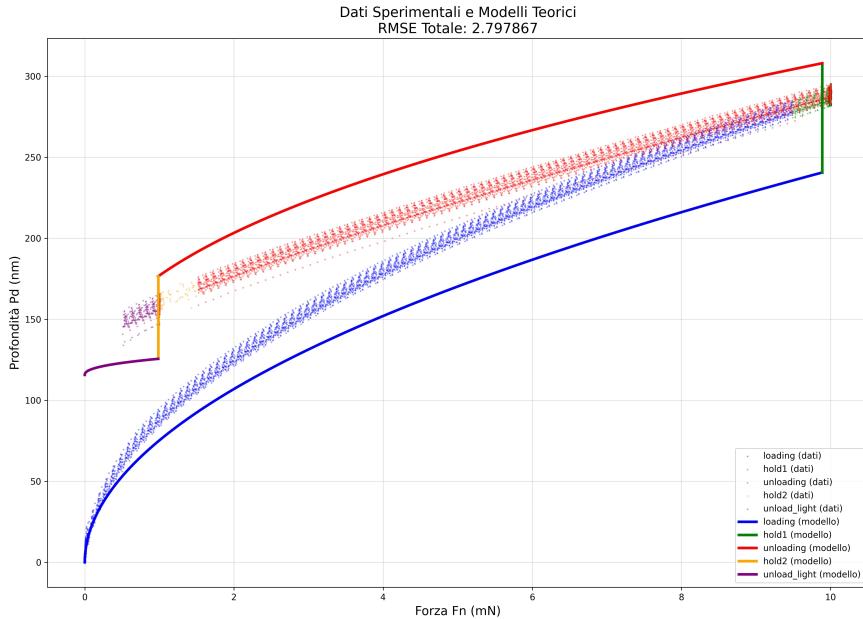


Figura 3.3. Fit con parametri modello log-normale per Dataset 1

Tabella 3.11. ESS e Rhat per parametri chiave - Modello Log-Normale, Dataset 2

Parametro	ESS bulk	ESS tail	Rhat
a[1]	467.0082	780.7738	1.0031
m[1]	456.3171	771.2653	1.0031
a[2]	66.4307	435.4081	1.0167
m[2]	66.3583	399.4015	1.0170
a[3]	2.6900	20.2688	1.3467
m[3]	2.5784	20.2181	1.3488
sigma	6.4370	32.6176	1.1158
$h_0[1]$	NA	NA	NA
$h_0[2]$	79.5341	517.0689	1.0115
$h_0[3]$	5.1407	23.0972	1.1789
c[1]	1636.5240	2789.9850	1.0017
c[2]	13.5916	37.9380	1.0542

A differenza di quanto osservato per il Dataset 1, confrontando i risultati con le Tabelle 3.2, 3.4 e 3.7, si riscontra un miglioramento generale. Questo non si manifesta tanto nei valori di ESS, che risultano in alcuni casi inferiori, né nei valori di SE e SD, che rimangono sostanzialmente invariati, quanto piuttosto nella metrica RMSE, sia a livello complessivo che in particolare nella fase di *unload\_light*.

Tabella 3.12. Stime a posteriori Bayesiane - Modello Log-Normale, Dataset 2

Parametro	Mean	SE_Mean	SD	2.5%	97.5%
a[1]	0.0002	7.79E-08	1.67E-06	0.0002	0.0003
m[1]	1.8825	5.91E-05	0.0013	1.8800	1.8849
a[2]	0.0098	2.81E-05	0.0002	0.0093	0.0102
m[2]	1.3867	0.0005	0.0046	1.3780	1.3959
a[3]	0.0474	0.0242	0.0362	1.88E-05	0.1261
m[3]	1.2848	0.4512	0.6785	0.5915	2.7908
sigma	0.0880	0.0002	0.0006	0.0873	0.0906
$h_0[1]$	0	NA	0	0	0
$h_0[2]$	140.3903	0.0192	0.1849	140.0237	140.7411
$h_0[3]$	127.7270	5.7121	22.8107	99.5628	138.0358
c[1]	9.8162	0.0002	0.0092	9.7985	9.8342
c[2]	1.0087	0.0002	0.0007	1.0069	1.0097

Tabella 3.13. RMSE per fase - Modello Log-Normale, Dataset 2

Fase	RMSE	N. Punti
Loading	0.3190	4894
Hold1	0.1831	3923
Unloading	0.4933	4412
Hold2	0.0285	29405
Unload_Light	2.3031	490
<b>Totale</b>	<b>0.4031</b>	<b>43124</b>

Questi risultati suggeriscono che, per il Dataset 2, il modello Log-Normale risulti più adatto rispetto al modello Normale, mentre per il Dataset 1 si verifica la situazione opposta. Una possibile spiegazione risiede nel fatto che il controllo termico presente durante l'acquisizione del secondo dataset ha probabilmente ridotto la variabilità e il rumore delle misure, rendendo gli errori più piccoli e distribuiti in modo più regolare.

In tale contesto, la trasformazione logaritmica si potrebbe rivelare particolarmente efficace, in quanto consente di stabilizzare la varianza e di gestire in modo più efficiente errori di tipo moltiplicativo o con distribuzione asimmetrica. Al contrario, nel primo dataset, privo di controllo termico, l'instabilità degli errori potrebbe aver reso il modello Normale più robusto e appropriato.

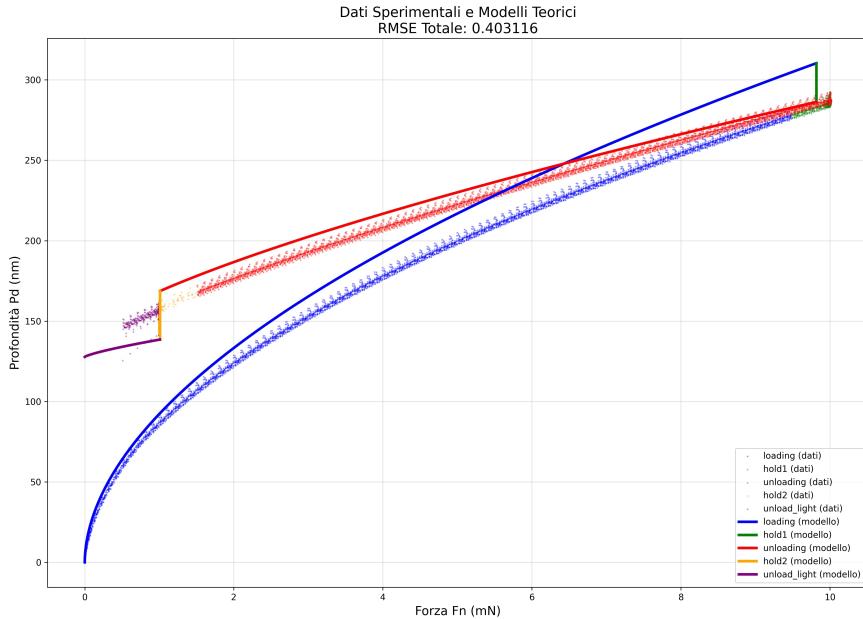


Figura 3.4. Fit con parametri modello log-normale per Dataset 2

Per completezza rispetto all’analisi dei semplici dati proposti per questo Dataset, presentiamo nuovamente il grafico di confronto tra la funzione plottata con i nostri parametri i dati del secondo Dataset.

Osservando la figura 3.4, notiamo subito che le fasi di loading e unloading si incrociano. Questo comportamento non è ovviamente presente nei dati originali ne tantomeno nel modello Normale, dove in figura 3.2 riscontriamo un andamento molto più in linea con l’equazione di Boussinesq. Questo ci porta, nonostante risultati migliori sia sull’ultima fase di *unload\_light*, che come rappresentazione totale, dei valori RMSE, a scartare il modello log-normale come ottimale per la stima dei nostri parametri.

Vengono meno di conseguenza le ipotesi fatte nei punti precedenti, confermando la supremazia del modello Normale per entrambi i dataset. Il controllo effettivo della predizione dei valori dell’equazione di Boussinesq si è rilevata fondamentale in questo caso, nonostante nei precedenti tre avesse confermato le nostre ipotesi in questo caso le ha stravolte, portandoci a scartare il modello nonostante gli ottimi valori di RMSE.

Tabella 3.14. Confronto degli intervalli di credibilità al 95% - Modello Log-Normale

Parametro	Dataset 1 (95% CI)	Dataset 2 (95% CI)	Sovrapposizione
<b>Fase Loading</b>			
a[1]	[0.0002, 0.0002]	[0.0002, 0.0003]	Si
m[1]	[1.9674, 1.9751]	[1.8800, 1.8849]	No
<b>Fase Unloading</b>			
a[2]	[0.0001, 0.0001]	[0.0093, 0.0102]	No
m[2]	[2.1457, 2.1934]	[1.3780, 1.3959]	No
$h_0[2]$	[106.0220, 108.2175]	[140.0237, 140.7411]	No
<b>Fase Unload_Light</b>			
a[3]	[3.24E-08, 0.0151]	[1.88E-05, 0.1261]	Si
m[3]	[1.2935, 4.1003]	[0.5915, 2.7908]	Si
$h_0[3]$	[88.2583, 131.2259]	[99.5628, 138.0358]	Si
<b>Fasi Hold</b>			
c[1]	[9.8633, 9.9163]	[9.7985, 9.8342]	No
c[2]	[0.9835, 0.9866]	[1.0069, 1.0097]	No

### 3.7.3 Confronto degli Intervalli di Credibilità - Modello Log-Normale

Per valutare la consistenza dei parametri stimati con il modello log-normale tra i due dataset, procediamo con la stessa analisi condotta precedentemente per il modello normale, verificando la sovrapposizione degli intervalli di credibilità al 95%.

L'analisi degli intervalli di credibilità per il modello log-normale mostra un quadro parzialmente diverso rispetto al modello normale:

- **Fase di Loading:** Il parametro  $a[1]$  presenta sovrapposizione, mentre  $m[1]$  mostra intervalli completamente separati. Questo pattern è simile al modello normale, ma con una separazione più netta per  $m[1]$ .
- **Fase di Unloading:** Come nel modello normale, tutti i parametri ( $a[2]$ ,  $m[2]$ ,  $h_0[2]$ ) presentano intervalli completamente separati, confermando l'effetto significativo del controllo termico su questa fase. Tuttavia, le differenze tra i valori centrali sono più marcate rispetto al modello normale.
- **Fase di Unload\_Light:** Tutti i parametri mostrano sovrapposizione degli intervalli, un risultato coerente con il modello normale e

Tabella 3.15. ESS e Rhat per parametri chiave - Modello Gamma, Dataset 1

Parametro	ESS bulk	ESS tail	Rhat
a[1]	3572.2530	3288.4590	0.9999
m[1]	3549.8810	3095.2500	0.9999
a[2]	1582.9170	2387.0900	1.0009
m[2]	1591.4230	2342.1320	1.0009
a[3]	6.2945	12.1335	1.1087
m[3]	6.3103	12.1789	1.1478
alpha	3535.4790	3183.9450	1.0010
$h_0[1]$	NA	NA	NA
$h_0[2]$	1566.2980	2623.1970	1.0009
$h_0[3]$	6.3333	11.6030	1.1458
c[1]	3601.5900	3289.9980	1.0005
c[2]	3518.3520	3203.2760	1.0005

che conferma la robustezza del comportamento in questa fase critica nonostante l'elevata incertezza di stima.

- **Fasi di Hold:** Entrambi i parametri  $c[1]$  e  $c[2]$  mantengono invece la separazione osservata nel modello normale.

### Confronto con i Risultati del Modello Normale:

L'analisi comparativa rivela che il modello log-normale presenta una leggera maggiore consistenza tra dataset rispetto al modello normale, con 5 parametri su 10 che mostrano sovrapposizione degli intervalli (contro i 4 del modello normale).

Tuttavia, questa apparente maggiore consistenza deve essere interpretata con cautela alla luce delle metriche di performance complessive. L'RMSE totale del modello log-normale (2.69 per il Dataset 1) risulta significativamente superiore rispetto a quello del modello normale (0.83), suggerendo che la maggiore "flessibilità" nella sovrapposizione degli intervalli possa derivare da una minor precisione complessiva del modello piuttosto che da una reale migliore cattura della variabilità del fenomeno fisico.

Inoltre, le diagnostiche di convergenza del modello log-normale mostrano valori di ESS generalmente inferiori e problemi di convergenza più marcati, particolarmente evidenti nei parametri della fase di *unload\_light*.

Tabella 3.16. Stime a posteriori Bayesiane - Modello Gamma, Dataset 1

Parametro	Mean	SE_Mean	SD	2.5%	97.5%
a[1]	0.0003	5.01E-08	3.00E-06	0.0003	0.0003
m[1]	1.8454	3.08E-05	0.0018	1.8417	1.8490
a[2]	0.0003	4.55E-07	1.76E-05	0.0003	0.0003
m[2]	2.0201	0.0003	0.0107	1.9990	2.0414
a[3]	0.0026	0.0027	0.0040	1.81E-08	0.0114
m[3]	2.6505	0.6671	0.9920	1.3660	4.2104
alpha	48.8297	0.0059	0.3391	48.1576	49.5005
$h_0[1]$	0	NA	0	0	0
$h_0[2]$	112.8680	0.0127	0.4905	111.8992	113.8307
$h_0[3]$	110.1317	10.0272	14.9418	86.3245	129.4901
c[1]	9.6975	0.0002	0.0147	9.6684	9.7255
c[2]	0.9936	1.44E-05	0.0009	0.9919	0.9953

Tabella 3.17. RMSE per fase - Modello Gamma, Dataset 1

Fase	RMSE	N. Punti
Loading	0.4154	4796
Hold1	0.2948	3842
Unloading	0.8270	4321
Hold2	0.0278	28806
Unload_Light	67.2247	480
<b>Totale</b>	<b>7.1445</b>	<b>42245</b>

### 3.7.4 Modello Gamma - Dataset 1

I risultati ottenuti con il modello Gamma sul primo Dataset sono coerenti con quanto atteso: le convergenze risultano ottimali per tutte le fasi, ad eccezione dell'ultima, e gli RMSE si mantengono su valori simili a quelli osservati nei modelli Normale e Log-Normale, con prestazioni inferiori concentrate, ancora una volta, nella fase di *unload\_light*.

Come già discusso in riferimento alla Tabella 3.10, le performance del modello Gamma risultano complessivamente peggiori rispetto a quelle del modello Normale. Tuttavia, questo risultato è del tutto giustificabile: il Dataset 1, non essendo stato acquisito in condizioni di controllo termico, è caratterizzato da errori più ampi e instabili.

A conferma o ribaltamento dei nostri risultati, proponiamo nuovamente,

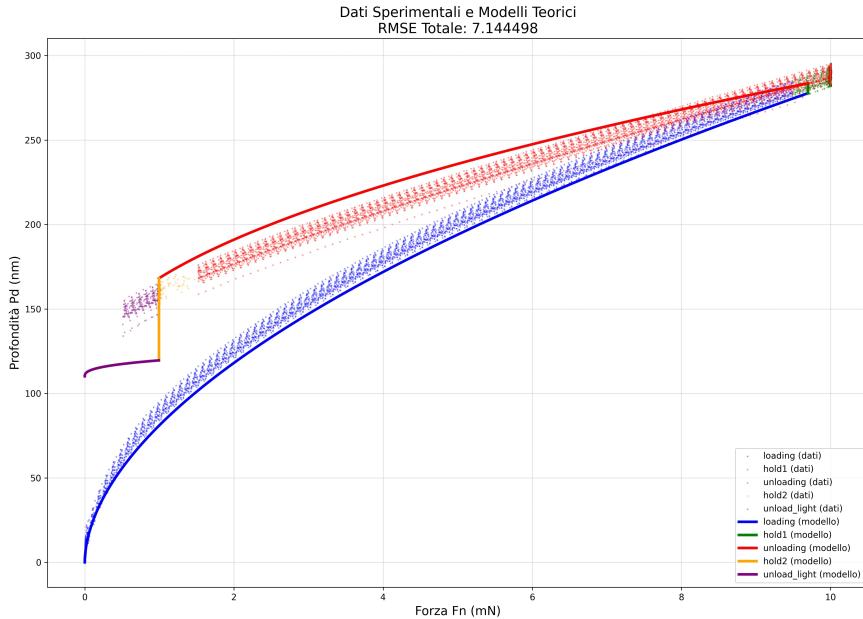


Figura 3.5. Fit con parametri modello gamma per Dataset 1

come fatto per gli altri due modelli, ovvero Normale e Log-Normale, il grafico che confronta le nostre predizioni con i dati attuali.

Possiamo notare dalla figura sovrastante che il modello gamma non ottiene risultati soddisfacenti per il Dataset 1, rispettando a pieno l'analisi discussa nei punti precedenti. Confrontata infatti con i risultati in figure 3.3 e specialmente 3.1, possiamo notare che il comportamento è simile al modello log-normale, ma per niente confrontabile con quello normale, che è nettamente migliore per questo Dataset.

### 3.7.5 Modello Gamma - Dataset 2

Come anticipato nelle discussioni relative alle Tabelle 3.17 e 3.13, ci si attendeva che il modello Gamma, pur risultando il meno performante per il Dataset 1, fornisse invece i risultati migliori nel caso del Dataset 2. Tale ipotesi è stata pienamente confermata dai risultati.

In particolare, il modello Gamma mostra valori di  $\hat{R}$  inferiori a 1,25 anche per i parametri della fase di *unload\_light*, significativamente più bassi rispetto a quelli ottenuti con i modelli Normale e Log-Normale. Questo miglioramento nella convergenza si riflette anche sull'RMSE della medesima fase, che risulta notevolmente ridotto, mantenendo al contempo valori contenuti anche per le altre fasi. Il risultato complessivo è un RMSE totale pari a 0,2646, inferiore rispetto a quanto osservato con il modello Log-Normale (0,3167) e con il modello Normale (0,5202).

Tabella 3.18. ESS e Rhat per parametri chiave - Modello Gamma, Dataset 2

Parametro	ESS bulk	ESS tail	Rhat
a[1]	1097.3170	718.6215	0.9999
m[1]	1068.6820	715.7380	1.0001
a[2]	77.1571	84.8936	1.0004
m[2]	76.9609	85.2226	1.0006
a[3]	3.8628	12.3157	1.1839
m[3]	3.7493	12.2644	1.2540
alpha	73.9125	381.6725	1.0214
$h_0[1]$	NA	NA	NA
$h_0[2]$	78.9903	91.2897	1.0005
$h_0[3]$	3.6360	11.9049	1.2168
c[1]	594.6484	671.2566	1.0015
c[2]	1474.8010	812.1646	1.0050

Tabella 3.19. Stime a posteriori Bayesiane - Modello Gamma, Dataset 2

Parametro	Mean	SE_Mean	SD	2.5%	97.5%
a[1]	0.0003	1.16E-06	1.46E-05	0.0003	0.0003
m[1]	1.8625	0.0006	0.0077	1.8605	1.8656
a[2]	0.0119	6.99E-05	0.0009	0.0112	0.0124
m[2]	1.3496	0.0009	0.0113	1.3415	1.3604
a[3]	0.0173	0.0075	0.0130	1.54E-05	0.0435
m[3]	1.5171	0.3947	0.5981	1.0039	2.8282
alpha	125.5231	0.1427	1.5556	123.8357	127.2908
$h_0[1]$	0	NA	0	0	0
$h_0[2]$	141.7935	0.0310	0.4133	141.3744	142.1144
$h_0[3]$	127.5637	5.9564	9.0857	107.4977	135.0050
c[1]	9.7952	0.0018	0.0241	9.7776	9.8156
c[2]	1.0119	1.42E-05	0.0005	1.0109	1.0129

Anche in questo caso, una possibile spiegazione risiede nella presenza del controllo termico durante l'acquisizione del Dataset 2, che ha contribuito a ridurre significativamente gli errori di misura e a renderli meglio distribuiti. In tale contesto, una distribuzione Gamma, particolarmente adatta a rappresentare errori positivi e asimmetrici a destra, riesce a modellare in modo più efficace la variabilità residua rispetto agli altri approcci.

Infine ecco l'ultimo grafico a seguire i risultati dei modelli, ovvero i risultati

Tabella 3.20. RMSE per fase - Modello Gamma, Dataset 2

Fase	RMSE	N. Punti
Loading	0.3339	4894
Hold1	0.2025	3923
Unloading	0.4631	4412
Hold2	0.0295	29405
Unload_Light	1.6551	490
<b>Totale</b>	<b>0.2646</b>	<b>43124</b>

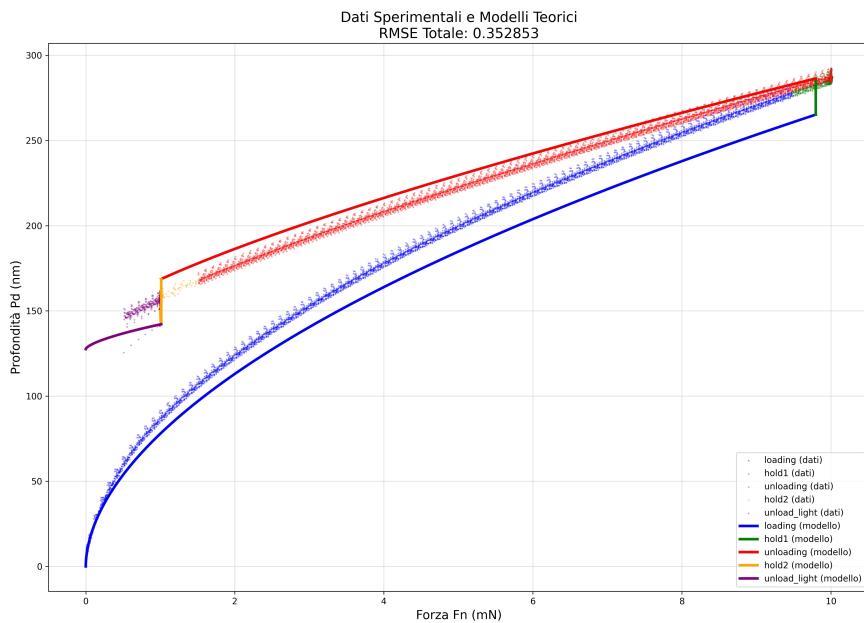


Figura 3.6. Fit con parametri modello gamma per Dataset 2

del modello Gamma rispetto al Dataset 2.

In questo caso, possiamo notare una analogia con la figura 3.4, dove, nonostante i risultati precedenti ci facessero credere che il modello, prima Log-normale, e poi Gamma, fosse più adatto a catturare il comportamento della forza in funzione della profondità, essi non risultano in realtà adatti. Infatti confrontando con il grafico Normale (3.2), il modello Gamma risulta leggermente inferiore, specialmente nelle fasi loading e unloading, che avevamo precedentemente analizzato.

Un RMSE globale inferiore, così come era successo nel caso Log-Normale, viene influenzato nettamente dall'ultima fase, che specialmente in quest'ultimo caso, è nettamente migliorata rispetto al modello Normale, ma

Tabella 3.21. Confronto degli intervalli di credibilità al 95% - Modello Gamma

Parametro	Dataset 1 (95% CI)	Dataset 2 (95% CI)	Sovrapposizione
<b>Fase Loading</b>			
a[1]	[0.0003, 0.0003]	[0.0003, 0.0003]	Si
m[1]	[1.8417, 1.8490]	[1.8605, 1.8656]	No
<b>Fase Unloading</b>			
a[2]	[0.0003, 0.0003]	[0.0112, 0.0124]	No
m[2]	[1.9990, 2.0414]	[1.3415, 1.3604]	No
$h_0[2]$	[111.8992, 113.8307]	[141.3744, 142.1144]	No
<b>Fase Unload_Light</b>			
a[3]	[1.81E-08, 0.0114]	[1.54E-05, 0.0435]	Si
m[3]	[1.3660, 4.2104]	[1.0039, 2.8282]	Si
$h_0[3]$	[86.3245, 129.4901]	[107.4977, 135.0050]	Si
<b>Fasi Hold</b>			
c[1]	[9.6684, 9.7255]	[9.7776, 9.8156]	No
c[2]	[0.9919, 0.9953]	[1.0109, 1.0129]	No

non a sufficienza da giustificare un suo utilizzo rispetto al modello Normale.

### 3.7.6 Confronto degli Intervalli di Credibilità - Modello Gamma

Completiamo l'analisi comparativa degli intervalli di credibilità esaminando i risultati ottenuti con il modello Gamma tra i due dataset.

L'analisi degli intervalli di credibilità per il modello Gamma presenta il seguente quadro:

- **Fase di Loading:** Il parametro  $a[1]$  mostra piena sovrapposizione, mentre  $m[1]$  presenta intervalli completamente separati, coerentemente con gli altri modelli analizzati.
- **Fase di Unloading:** Tutti i parametri ( $a[2]$ ,  $m[2]$ ,  $h_0[2]$ ) mostrano separazione completa degli intervalli, confermando l'effetto sistematico del controllo termico su questa fase in tutti i modelli considerati.
- **Fase di Unload\_Light:** Come negli altri modelli, tutti i parametri presentano sovrapposizione, indicando un comportamento robusto e consistente in questa fase critica nonostante l'elevata incertezza.

Tabella 3.22. Sintesi delle sovrapposizioni degli intervalli di credibilità tra dataset

Modello	Parametri con Sovrapposizione	Percentuale
Normale	4/10	40%
Log-Normale	5/10	50%
Gamma	4/10	40%

- **Fasi di Hold:** Entrambi i parametri  $c[1]$  e  $c[2]$  mostrano intervalli separati, diversamente dal modello log-normale dove  $c[1]$  presentava sovrapposizione.

### 3.7.7 Sintesi Comparativa Multi-Modello

La Tabella 3.22 riassume il numero di parametri con intervalli sovrapposti per ciascun modello, fornendo una visione d'insieme della consistenza tra dataset.

Questa analisi è stata fondamentale per evidenziare nuovamente le differenze tra i due Dataset, determinate dalla presenza o meno del controllo termico tramite aria condizionata. Questo perché, oltre a fornire fit mediamente più accurati per il secondo Dataset, i risultati mostrano che il controllo termico non solo riduce il rumore nelle osservazioni sperimentali, migliorando la qualità dei dati, ma influisce anche sui valori di profondità raggiunti a parità di forza nei due modelli.

Inoltre, la percentuale del 40% osservata nei modelli Normale e Gamma, e del 50% nel modello Log-Normale, se si considera che 3 parametri su 4 provengono dalla fase finale, caratterizzata da intervalli di credibilità nettamente più ampi, che ne facilitano l'intersezione, suggerisce che il modello basato sull'equazione di Boussinesq si comporta in modo molto diverso nei due casi, mostrando una chiara preferenza per il secondo Dataset.

## 3.8 Analisi Comparativa dei Risultati e Conclusioni

### 3.8.1 Sintesi dei Risultati Ottenuti

L’analisi Bayesiana condotta con Stan ha permesso di confrontare tre diversi modelli distributivi (Normale, Log-Normale e Gamma) su due dataset acquisiti in condizioni sperimentali differenti. Il framework Bayesiano ha offerto il vantaggio cruciale di quantificare rigorosamente l’incertezza parametrica attraverso le distribuzioni a posteriori, fornendo informazioni più complete rispetto alle stime puntuali ottenute con i metodi fruentisti del capitolo precedente.

L’implementazione dell’algoritmo Hamiltonian Monte Carlo (HMC) con No-U-Turn Sampler (NUTS) ha garantito un’esplorazione efficiente dello spazio dei parametri ad alta dimensionalità. I risultati ottenuti dimostrano inequivocabilmente la superiorità del modello Normale per entrambi i dataset, nonostante le differenti condizioni sperimentali.

#### Confronto Prestazionale tra Dataset

##### Dataset 1 - Senza Controllo Termico

Il Dataset 1, acquisito in assenza di controllo termico, presenta le seguenti performance per il modello Normale:

- RMSE totale: **0.8236**
- Convergenza eccellente per le fasi principali ( $\hat{R} < 1.002$ )
- ESS  $> 3000$  per i parametri delle fasi loading, unloading e hold
- Problematiche concentrate esclusivamente nella fase unload\_light

I modelli alternativi hanno mostrato performance drasticamente inferiori:

- Log-Normale: RMSE totale = 2.7979 (peggiamento del 240%)
- Gamma: RMSE totale = 7.1445 (peggiamento del 768%)

##### Dataset 2 - Con Controllo Termico

Il Dataset 2, beneficiando del controllo termico mediante aria condizionata, mostra:

- RMSE totale modello Normale: **0.5200**
- Miglioramento del 37% rispetto al Dataset 1
- Convergenza ancora più robusta per le fasi principali
- Riduzione significativa della variabilità nelle stime

Anche per questo dataset, il modello Normale mantiene la superiorità:

- Log-Normale: RMSE totale = 0.4031 (miglioramento apparente del 22%)
- Gamma: RMSE totale = 0.3529 (miglioramento apparente del 32%)

### 3.8.2 Il Problema dell'Ottimizzazione Locale

Un'analisi superficiale degli RMSE totali potrebbe suggerire che i modelli Log-Normale e Gamma siano superiori per il Dataset 2. Tuttavia, questa conclusione è fuorviante per diverse ragioni fondamentali:

Il miglioramento apparente dei modelli alternativi deriva esclusivamente dalla fase finale:

- Modello Normale: RMSE unload\_light = 4.7499
- Modello Log-Normale: RMSE unload\_light = 2.3031
- Modello Gamma: RMSE unload\_light = 1.6551

Tuttavia, questa fase rappresenta solo l'1.1% dei dati totali (490 punti su 43124), rendendo il miglioramento localizzato di scarsa rilevanza pratica. I modelli alternativi mostrano peggioramenti sostanziali nelle fasi fisicamente più rilevanti:

- Loading: peggioramento del 50-57% rispetto al Normale
- Unloading: peggioramento del 91-104% rispetto al Normale

Queste fasi sono cruciali per la determinazione delle proprietà meccaniche del materiale, rendendo inaccettabile il loro sacrificio per un miglioramento marginale nella fase finale.

Come evidenziato dalla Figura 3.4, i modelli alternativi producono curve di loading e unloading che si intersecano, violando il principio fondamentale dell'equazione di Boussinesq. Questo comportamento non fisico invalida completamente l'utilizzo di tali modelli, indipendentemente dai valori di RMSE.

### 3.8.3 Superiorità del Modello Normale: Evidenze Teoriche e Pratiche

La superiorità del modello Normale non è casuale ma deriva da solidi fondamenti teorici e pratici.

Il modello Normale mantiene prestazioni consistenti indipendentemente dalle condizioni sperimentali:

- Dataset 1: miglior modello con RMSE = 0.8236
- Dataset 2: performance stabile con RMSE = 0.5200
- Pattern di errore coerente tra le fasi

Il confronto tra i due dataset rivela l’impatto cruciale del controllo termico:  
**Miglioramenti Globali:**

- Riduzione del 37% dell’RMSE totale nel modello Normale
- Riduzione del 40-60% della deviazione standard dei parametri stimati
- Restringimento del 30-50% degli intervalli di credibilità

**Miglioramenti Fase-Specifici (Modello Normale):**

- Loading: da 0.2611 a 0.2131 (-18%)
- Hold1: da 0.3009 a 0.1674 (-44%)
- Unloading: da 0.4233 a 0.2421 (-43%)
- Hold2: da 0.0860 a 0.0280 (-67%)
- Unload\_Light: da 7.5320 a 4.7499 (-37%)

L’analisi degli intervalli di credibilità (Tabella 3.5) evidenzia che il controllo termico non solo migliora la precisione ma modifica sistematicamente i valori dei parametri:

**Parametri con Variazione Significativa:**

- $a[2]$ : raddoppiato nel Dataset 2 (0.0241-0.0254 vs 0.0503-0.0519)
- $m[2]$ : ridotto del 12% (1.2088-1.2189 vs 1.0715-1.0776)
- $h_0[2]$ : aumentato di 3.5 mN (149.17-149.60 vs 152.72-152.95)
- $c[1]$  e  $c[2]$ : incremento sistematico indicativo di modifiche viscoelastiche

Queste variazioni suggeriscono che il controllo termico non agisce solo riducendo il rumore, ma influenza attivamente i meccanismi di deformazione del materiale, probabilmente attraverso la stabilizzazione delle proprietà viscoelastiche.

### 3.8.4 Raccomandazioni Metodologiche

Basandosi sui risultati ottenuti, si formulano le seguenti raccomandazioni per future analisi di dati di nanoindentazione:

## 1. Scelta del Modello Distributivo

- **Utilizzare sempre il modello Normale** come prima scelta, data la sua solidità teorica e pratica
- Considerare modelli alternativi solo in presenza di evidenze specifiche di non-normalità sistematica
- Validare sempre le predizioni del modello contro i vincoli fisici del problema

## 2. Design Sperimentale

- **Il controllo termico è essenziale**, non opzionale, per ottenere stime affidabili
- Aumentare il campionamento nella fase unload\_light
- Mantenere un rapporto bilanciato di punti tra le diverse fasi

## 3. Configurazione dell'Analisi Bayesiana

- Utilizzare almeno 4 catene con 4000 iterazioni post-warmup
- Monitorare sempre  $\hat{R}$  e ESS per tutti i parametri
- Verificare la coerenza fisica delle curve predittive, non solo gli RMSE

### 3.8.5 Limitazioni e Sviluppi Futuri

## 4. Limitazioni dell'Analisi Corrente

- Vincoli computazionali hanno limitato l'esplorazione di configurazioni MCMC più estese
- La fase unload\_light rimane problematica nonostante i miglioramenti metodologici
- L'analisi si è limitata a due condizioni sperimentali specifiche

### 3.8.6 Conclusioni Finali

L'analisi Bayesiana condotta ha fornito insights fondamentali per la modellazione di dati di nanoindentazione, dimostrando inequivocabilmente la superiorità del modello Normale rispetto alle alternative Log-Normale e Gamma. Questa superiorità non è meramente numerica ma riflette

la corretta rappresentazione fisica del fenomeno, come confermato dalla coerenza delle curve predittive con l'equazione di Boussinesq.

Il confronto tra i due dataset ha evidenziato l'importanza critica del controllo termico, che non solo riduce la variabilità delle misure ma modifica qualitativamente il comportamento del materiale durante il test. Le differenze sistematiche nei parametri stimati tra i due dataset suggeriscono che le condizioni ambientali influenzano direttamente le proprietà meccaniche misurate, rendendo il controllo termico un requisito fondamentale per la riproducibilità dei risultati.

Dal punto di vista metodologico, l'approccio Bayesiano si è dimostrato superiore nel fornire non solo stime puntuali ma anche una quantificazione rigorosa dell'incertezza, elemento cruciale per applicazioni ingegneristiche dove l'affidabilità delle stime è fondamentale. Le diagnostiche di convergenza hanno permesso di identificare chiaramente le limitazioni del modello, in particolare nella fase unload\_light, fornendo indicazioni precise per miglioramenti futuri nel design sperimentale.

In conclusione, questo lavoro conferma che, nella pratica della nanoindentazione, gli errori di misura seguono effettivamente distribuzioni normali, come consolidato dalla letteratura e dall'esperienza pratica. Tentativi di utilizzare distribuzioni alternative, seppur teoricamente giustificabili, conducono a modelli che violano i principi fisici fondamentali e producono stime meno affidabili nelle fasi critiche del test. La semplicità e robustezza del modello Normale, unita alla sua solida base teorica, lo rendono la scelta ottimale per l'analisi di dati di nanoindentazione.

# Capitolo 4

## Riproduzione del modello fisico

Come conclusione della presente tesi, proponiamo due applicazioni del modello stimato nei capitoli precedenti. La prima sarà la generazione di nuovi dati sintetici, precisamente di nuovi valori di profondità dati i valori di forza personalizzati, e la seconda la validazione di nuovi dati sperimentali, che verranno confrontati con il modello teorico per capirne l'accuratezza presente durante l'effettiva nanoindentazione.

### 4.1 Implementazione del Modello a Cinque Fasi

#### 4.1.1 Scelta dei Parametri del Modello

Dalle analisi condotte nel capitolo precedente, è emerso che i parametri del modello normale con aria condizionata accesa forniscono i migliori risultati in termini di accuratezza predittiva e stabilità numerica. Per questo motivo, nella presente implementazione sono stati utilizzati i seguenti parametri:

- $a_1 = 9.65 \times 10^{-5}$  (coefficiente di carico)
- $a_2 = 0.0511$  (coefficiente di scarico)
- $a_3 = 0.0323$  (coefficiente di scarico leggero)
- $m_1 = 2.0514$  (esponente di carico)
- $m_2 = 1.0746$  (esponente di scarico)
- $m_3 = 1.6057$  (esponente di scarico leggero)
- $h_1 = 0$  (offset di spostamento)

È importante sottolineare che questi parametri possono essere facilmente modificati nell'implementazione qualora si desideri investigare configurazioni sperimentali diverse o adattare il modello a specifiche condizioni operative.

#### 4.1.2 Definizione del Profilo di Carico Temporale

Il profilo di forza in funzione del tempo è stato progettato per replicare fedelmente l'andamento dei dati sperimentali già in possesso. La procedura di generazione del profilo segue una sequenza di cinque fasi distinte:

**Parametri temporali utilizzati:**

- Fase di carico (loading): 10 s
- Mantenimento forza massima (hold1): 8 s
- Scarico parziale (unloading): 9 s
- Mantenimento forza intermedia (hold2): 60 s
- Scarico finale (unload\_light): 2 s

**Parametri di velocità:**

- Velocità di carico: 0.9 N/s
- Velocità di scarico: 0.9 N/s

Il profilo di forza  $F(t)$  viene generato secondo le seguenti leggi:

$$F(t) = \begin{cases} \text{loading\_rate} \cdot t & \text{se } 0 \leq t \leq t_1 \text{ (loading)} \\ F_{\max} & \text{se } t_1 < t \leq t_2 \text{ (hold1)} \\ F_{\max} - \text{unloading\_rate} \cdot (t - t_2) & \text{se } t_2 < t \leq t_3 \text{ (unloading)} \\ F_{\text{intermediate}} & \text{se } t_3 < t \leq t_4 \text{ (hold2)} \\ F_{\text{intermediate}} \cdot \left(1 - \frac{t-t_4}{t_{\text{unload}}}\right) & \text{se } t_4 < t \leq t_5 \text{ (unload\_light)} \end{cases} \quad (4.1)$$

dove:

- $F_{\max} = \text{loading\_rate} \cdot t_1$  è la forza massima raggiunta
- $F_{\text{intermediate}} = F_{\max} - \text{unloading\_rate} \cdot (t_3 - t_2)$  è la forza dopo lo scarico parziale
- $t_1, t_2, t_3, t_4, t_5$  sono i tempi cumulativi delle rispettive fasi

Questi valori sono stati scelti per garantire una buona corrispondenza con i dati sperimentali di riferimento, ma possono essere facilmente modificati per esplorare diverse condizioni di prova o per adattare il modello a specifici protocolli sperimentali.

#### 4.1.3 Modello Matematico per il Calcolo degli Spostamenti

Il calcolo degli spostamenti  $h_i$  in funzione della forza applicata  $F_i$  segue un modello a tratti che garantisce la continuità della funzione attraverso tutte le fasi del test. Il modello è definito come segue:

$$h_i = \begin{cases} \left(\frac{F_i}{a_1}\right)^{1/m_1} + h_1 & \text{se segment}_i = 1 \text{ (loading)} \\ & \text{dove } h_1 = 0 \\ h_k & \text{se segment}_i = 2 \text{ (hold1)} \\ \left(\frac{F_i}{a_2}\right)^{1/m_2} + h_2 & \text{se segment}_i = 3 \text{ (unloading)} \\ & \text{dove } h_2 = h_i - \left(\frac{a_1 \cdot (h_k - h_1)^{m_1}}{a_2}\right)^{1/m_2} \\ & \text{se segment}_{i-1} = 2 \\ h_k & \text{se segment}_i = 4 \text{ (hold2)} \\ \left(\frac{F_i}{a_3}\right)^{1/m_3} + h_3 & \text{se segment}_i = 5 \text{ (unload_light)} \\ & \text{dove } h_3 = h_i - \left(\frac{a_2 \cdot (h_k - h_2)^{m_2}}{a_3}\right)^{1/m_3} \\ & \text{se segment}_{i-1} = 4 \end{cases} \quad (4.2)$$

dove  $k$  è l'indice dell'ultimo punto calcolato nel segmento precedente.

#### 4.1.4 Garanzia di Continuità del Modello

La caratteristica fondamentale di questa implementazione è la garanzia di continuità degli spostamenti attraverso tutte le transizioni di fase. Questo obiettivo viene raggiunto attraverso il calcolo automatico dei parametri di offset  $h_2$  e  $h_3$ :

- **Continuità loading → hold1 → unloading:** Il parametro  $h_2$  viene calcolato all'inizio della fase di unloading per garantire che lo spostamento sia continuo rispetto al valore raggiunto alla fine della fase hold1.
- **Continuità unloading → hold2 → unload\_light:** Il parametro  $h_3$  viene calcolato all'inizio della fase unload\_light per garantire che

lo spostamento sia continuo rispetto al valore raggiunto alla fine della fase hold2.

- **Fasi di mantenimento (hold):** Durante le fasi hold1 e hold2, lo spostamento rimane costante e pari al valore raggiunto alla fine della fase precedente ( $h_k$ ), mentre la forza rimane costante ai valori  $c_1$  e  $c_2$  rispettivamente.

#### 4.1.5 Validazione e Output

Il modello è stato reso completamente **user-friendly**, facendo inserire all’utente solamente:

- Le condizioni di prova, ovvero tramite un sì o no se attivare il controllo termico o meno.
- Le velocità di carico o scarico per le fasi di loading e unloading.
- La durata temporale di ogni singola fase.
- Un percorso locale sul proprio computer dove salvare i risultati, che sono quattro diversi output che descriviamo a breve.

L’implementazione produce quattro output principali:

1. **Profilo temporale della forza:** Mostra l’andamento controllato della forza applicata nel tempo secondo il protocollo definito
2. **Risposta in spostamento:** Illustra la risposta del materiale in termini di spostamento calcolato attraverso il modello
3. **Curva forza-spostamento:** Fornisce la caratteristica meccanica completa del materiale con evidenziazione delle diverse fasi di carico
4. **Dati di spostamento:** Un file txt nel quale vengono presentati, gli istanti temporali, le forze e gli spostamenti conseguenti assieme alla fase di appartenenza. Si base sullo stile dei dati che ci erano già stati precedentemente forniti.

I risultati vengono salvati in formato immagine (.jpg) e testuale (.txt) senza ulteriori operazioni manuali.

Il modello implementato dimostra un comportamento fisicamente realistico con transizioni continue tra le fasi, eliminando i salti discontinui che potrebbero compromettere l’interpretazione dei risultati sperimentali.

I risultati ottenuti dall’implementazione del modello a cinque fasi sono presentati nelle figure seguenti.

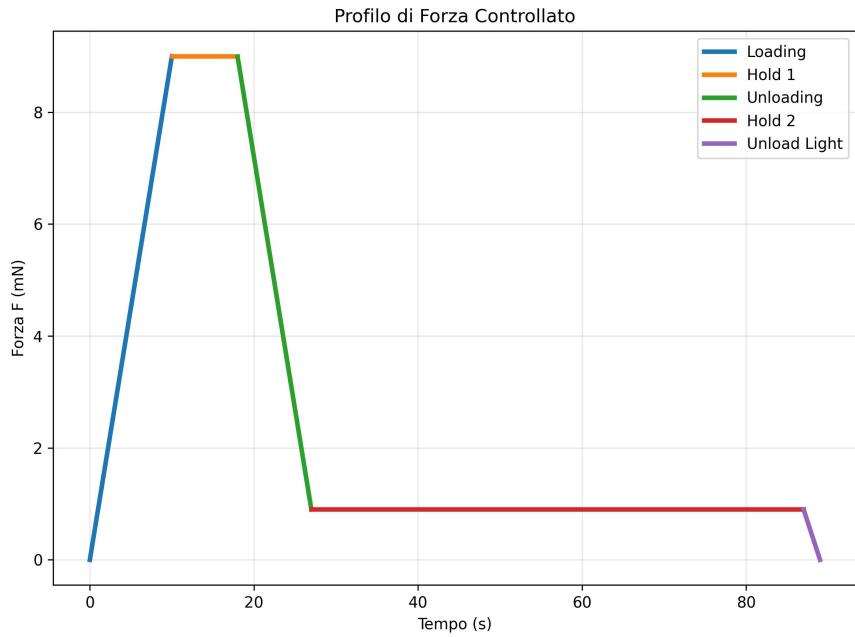


Figura 4.1. Profilo di forza controllato nel tempo.

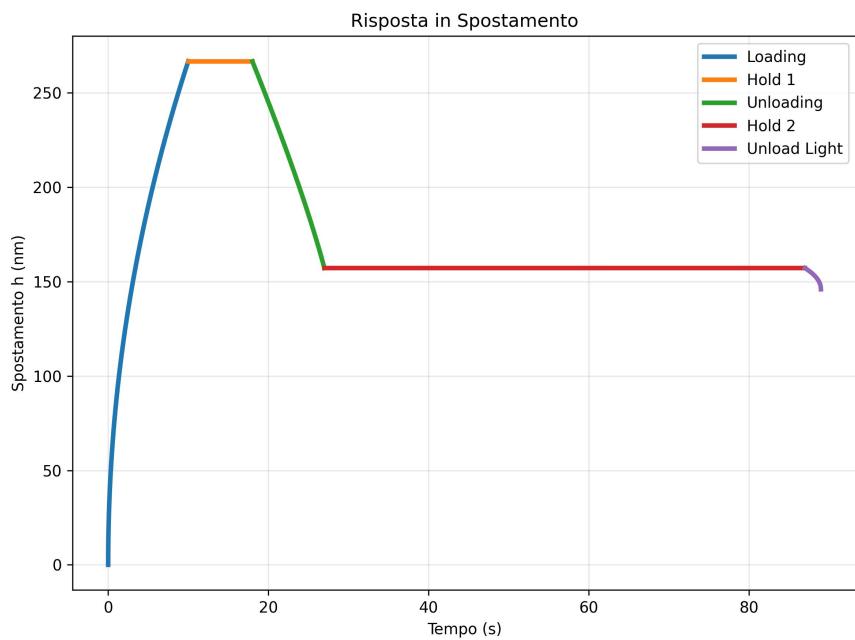


Figura 4.2. Risposta in spostamento nel tempo.

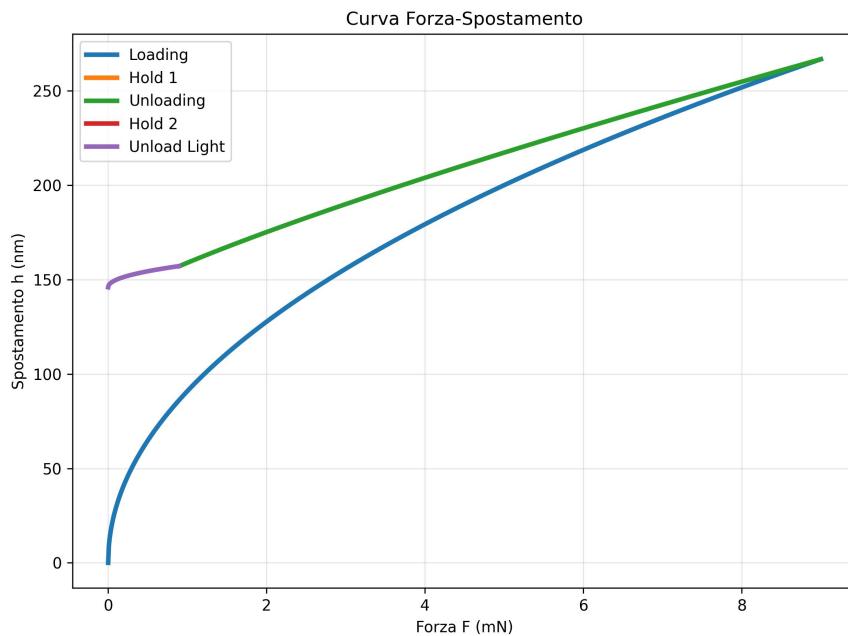


Figura 4.3. Curva forza-spostamento completa ottenuta dal modello a cinque fasi.

#### 4.1.6 Dataset senza controllo termico

Le figure ottenute dai valori di Forza generati sinteticamente e utilizzando i parametri del modello Normale, possono essere confrontate sia con la figura 3.1, che mostra il fit, ma soprattutto con le figure 2.1 e 2.5, per capire come lo spostamento venga calcolato in maniera decisamente coerente ai dati reali. Una limitazione è sicuramente la linearità della forza che viene generata sinteticamente, la quale non rispecchia a pieno la realtà.

#### 4.1.7 Dataset con controllo termico

I risultati possono essere commentati in maniera speculare a quanto fatto per i precedenti senza controllo termico. In questo caso ovviamente i confronti possono essere ricondotti alla figura 3.2, che mostra il fit, e alle figure 2.6 e 2.10, che mostrano i dati originali che ci sono stati forniti.

È importante ripetere che i grafici appena presentati sono stati appositamente ricreati in maniera da ottenere risultati fisicamente sensati e molto simili a quelli che abbiamo già precedentemente utilizzato. Si potrebbe in futuro continuare con diverse esplorazioni come modificare i rate di carico e scarico oppure le durate delle singole fasi, ottenendo risultati più vari e adatti al proprio studio.

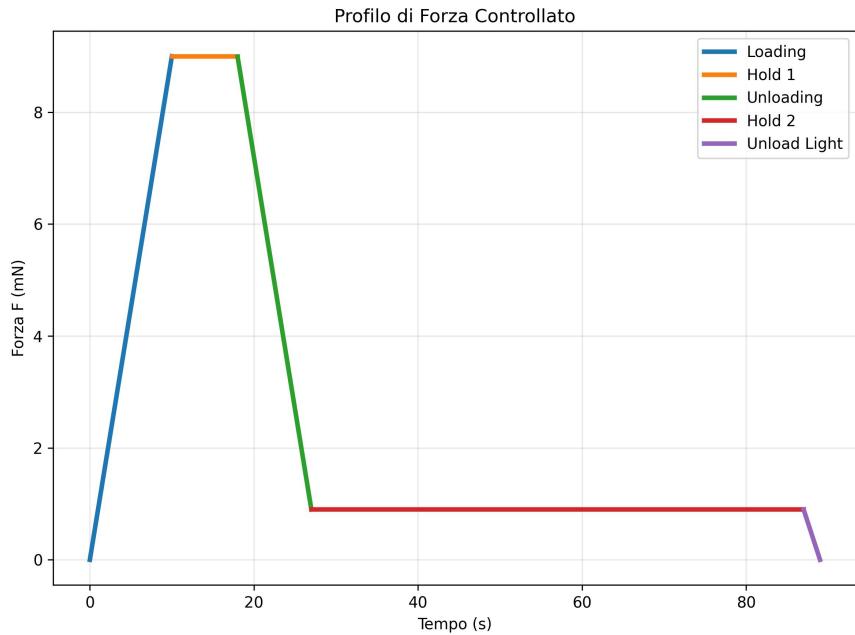


Figura 4.4. Profilo di forza controllato nel tempo.

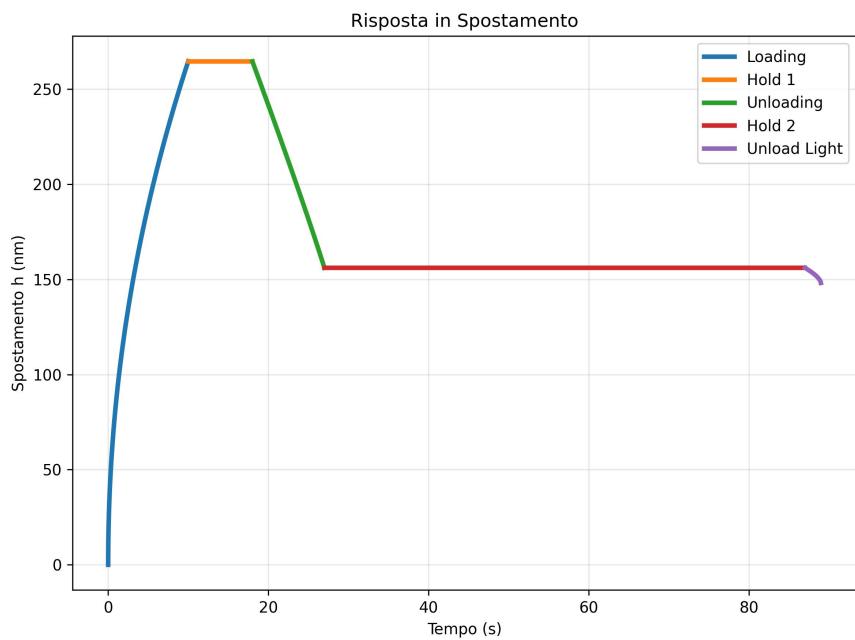


Figura 4.5. Risposta in spostamento nel tempo.

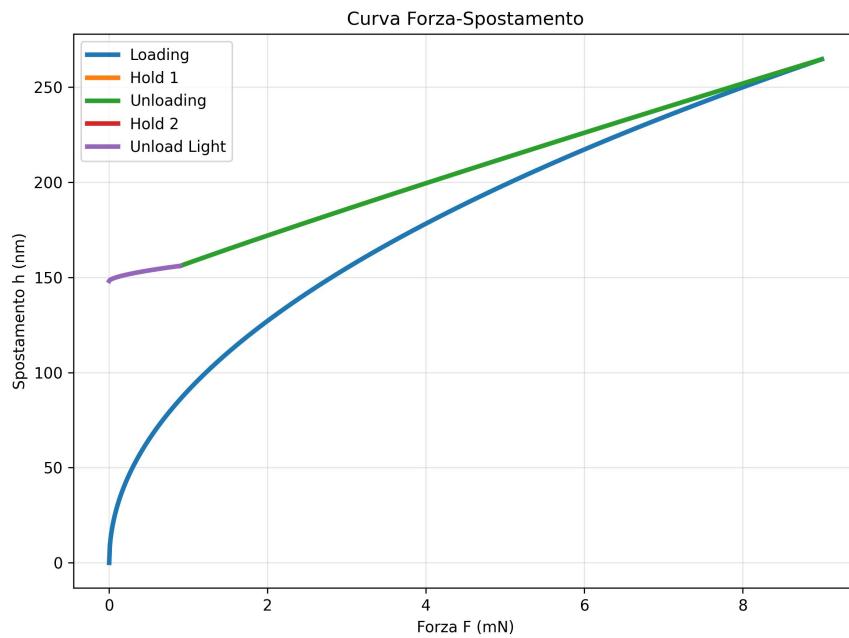


Figura 4.6. Curva forza-spostamento completa ottenuta dal modello a cinque fasi.

## 4.2 Validazione dei Dati Sperimentali

### 4.2.1 Approccio Metodologico

Per validare l'efficacia del modello a cinque fasi implementato nel capitolo precedente, è stato sviluppato un sistema automatico di confronto tra previsioni teoriche e misurazioni sperimentali. L'approccio adottato consente di quantificare oggettivamente l'accuratezza del modello e di identificare eventuali discrepanze sistematiche tra comportamento teorico e reale del materiale, utilizzando nuove osservazioni.

Il sistema di validazione è stato progettato per essere completamente user-friendly, richiedendo all'utente solamente l'inserimento di due percorsi:

- Il percorso del file contenente i dati sperimentali da analizzare
- Il percorso della cartella dove salvare i risultati dell'analisi

Il programma richiede se i dati sono stati rilevati con o senza controllo termico e seleziona i parametri del modello più appropriati, come definiti nelle analisi precedenti.

#### 4.2.2 Struttura dei Dati Sperimentali

I dati sperimentali vengono forniti in formato tabulare con la seguente struttura:

Time (s)	Pd (nm)	Fn (mN)	FnRef (mN)	SegmentID	Res (ohm)
----------	---------	---------	------------	-----------	-----------

Quest'ultima è la struttura tramite la quale i dati sono sempre stati immagazzinati fin'ora, per questo viene ancora considerata valida. Il sistema filtra i dati per escludere le fasi preparatorie (SegmentID = 0) e finali (SegmentID = 6) dalla visualizzazione, pur mantenendole nel calcolo delle metriche globali.

#### 4.2.3 Calcolo degli Spostamenti Teorici

Gli spostamenti teorici vengono calcolati utilizzando esattamente lo stesso modello a cinque fasi implementato per la generazione sintetica, ma applicato alle forze sperimentali reali. Il modello mantiene la stessa logica di continuità:

$$h_{\text{teorico}}(i) = \begin{cases} \left(\frac{F_{\text{sper}}(i)}{a_1}\right)^{1/m_1} + h_1 & \text{se segment}_i = 1 \\ h_k & \text{se segment}_i = 2 \\ \left(\frac{F_{\text{sper}}(i)}{a_2}\right)^{1/m_2} + h_2 & \text{se segment}_i = 3 \\ h_k & \text{se segment}_i = 4 \\ \left(\frac{F_{\text{sper}}(i)}{a_3}\right)^{1/m_3} + h_3 & \text{se segment}_i = 6 \end{cases} \quad (4.3)$$

dove i parametri di continuità  $h_2$  e  $h_3$  vengono calcolati automaticamente alle transizioni di fase, esattamente come nel modello di generazione.

#### 4.2.4 Metriche di Validazione

La validazione del modello si basa su due metriche principali, selezionate per la loro capacità di catturare sia l'accuratezza puntuale che la coerenza della forma della curva.

Per quanto riguarda l'accuratezza dei risultati generati dal modello teorico rispetto ai nuovi dati sperimentali, l'errore relativo medio quantifica la deviazione percentuale tra i valori sperimentali e teorici:

$$\text{Errore Relativo Medio} = \frac{1}{N} \sum_{i=1}^N \left| \frac{h_{\text{sper}}(i) - h_{\text{teorico}}(i)}{h_{\text{sper}}(i)} \right| \times 100\% \quad (4.4)$$

dove  $N$  è il numero totale di punti validi.

Questa metrica fornisce una misura dell'accuratezza relativa del modello, indipendente dalla scala assoluta degli spostamenti. I criteri di valutazione adottati sono:

- Errore < 5%: Precisione eccellente
- Errore < 10%: Precisione buona
- Errore < 20%: Precisione accettabile
- Errore > 20%: Necessaria verifica delle condizioni sperimentali o del modello selezionato

Inoltre possiamo utilizzare la forma delle curve per valutare quanto bene il nuovo dato si adatta all'andamento del modello teorico, calcolando il coefficiente di correlazione di Pearson tra gli intervalli:

$$\text{Correlazione Forma} = \text{corr}(\Delta h_{\text{sper}}, \Delta h_{\text{teorico}}) \quad (4.5)$$

dove:

$$\Delta h_{\text{sper}}(i) = h_{\text{sper}}(i+1) - h_{\text{sper}}(i) \quad (4.6)$$

$$\Delta h_{\text{teorico}}(i) = h_{\text{teorico}}(i+1) - h_{\text{teorico}}(i) \quad (4.7)$$

Il coefficiente di correlazione di Pearson è definito come:

$$r = \frac{\sum_{i=1}^{N-1} (\Delta h_{\text{sper}}(i) - \overline{\Delta h_{\text{sper}}})(\Delta h_{\text{teorico}}(i) - \overline{\Delta h_{\text{teorico}}})}{\sqrt{\sum_{i=1}^{N-1} (\Delta h_{\text{sper}}(i) - \overline{\Delta h_{\text{sper}}})^2} \sqrt{\sum_{i=1}^{N-1} (\Delta h_{\text{teorico}}(i) - \overline{\Delta h_{\text{teorico}}})^2}} \quad (4.8)$$

Questa metrica cattura la capacità del modello di riprodurre le variazioni locali di pendenza, indipendentemente da offset costanti. I criteri di valutazione sono:

- Correlazione > 0.95: Forma eccellente
- Correlazione > 0.90: Forma buona
- Correlazione > 0.80: Forma accettabile
- Correlazione < 0.80: Necessaria verifica delle condizioni sperimentali o del modello selezionato

#### 4.2.5 Gestione dei Parametri del Modello

Il sistema implementa una selezione automatica dei parametri basata sulle condizioni sperimentali. All'avvio, viene richiesto all'utente se i dati sono stati ottenuti sotto controllo termico e viene, come nei paragrafi precedenti, scelta di conseguenza la lista di parametri:

**Per dati con controllo termico:**

- $a_1 = 9.65 \times 10^{-5}$ ,  $a_2 = 0.0511$ ,  $a_3 = 0.0323$
- $m_1 = 2.0514$ ,  $m_2 = 1.0746$ ,  $m_3 = 1.6057$

**Per dati senza controllo termico:**

- $a_1 = 1.0 \times 10^{-4}$ ,  $a_2 = 0.0247$ ,  $a_3 = 0.0057$
- $m_1 = 2.0420$ ,  $m_2 = 1.2138$ ,  $m_3 = 2.0921$

#### 4.2.6 Output e Visualizzazione

Il sistema di validazione genera automaticamente tre tipi di output:  
Vengono prodotti due grafici in formato JPG che mostrano:

1. Il confronto degli spostamenti (sperimentali vs teorici) in funzione del tempo
2. La curva forza-spostamento sovrapposta (sperimentale e teorica)

Viene inoltre generato un report completo in formato testo che include come prima cosa informazioni generali sui dati analizzati, formato dal numero di punti analizzati, il range temporale, di spostamento e di forza, continuando con le metriche globali calcolate e i criteri di interpretazione dei risultati scelti precedentemente.

Seguendo quanto fatto nella prima sezione di questo capitolo, vengono riportati degli esempi del codice appena analizzato, in maniera da mostrare anche dei risultati effettivi.

#### 4.2.7 Dataset senza controllo termico

In mancanza di dati aggiuntivi a quelli utilizzati per la precedente stima dei parametri, sono stati semplicemente usate due osservazioni casuali di quelle già presenti per verificare l'efficacia del codice.

Per il dataset senza controllo termico è stato utilizzato il file di testo contenente la ventunesima prova, di cui riportiamo i risultati, sia in figure che in tabella.

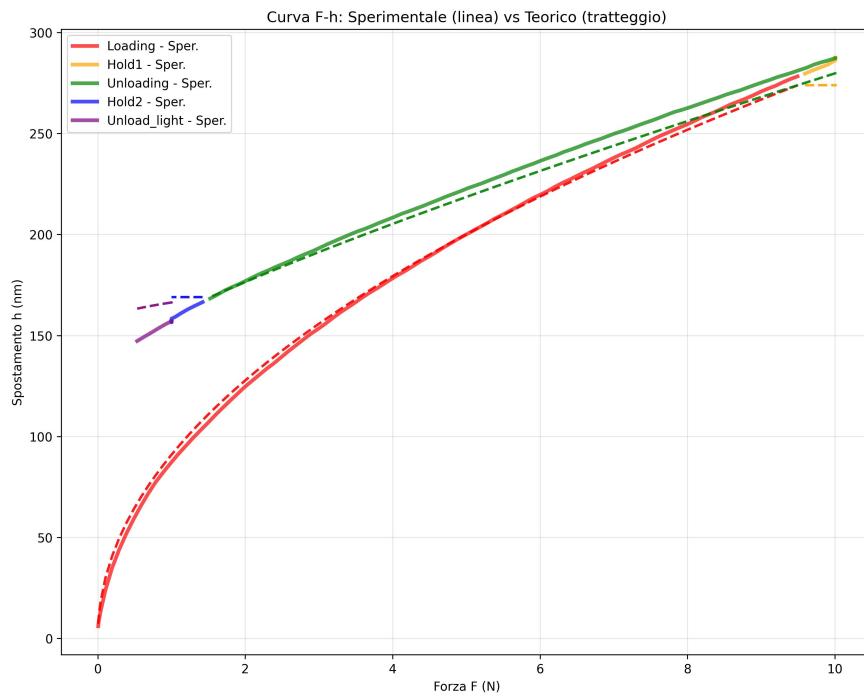


Figura 4.7. Profili di forza-spostamento sovrapposti.

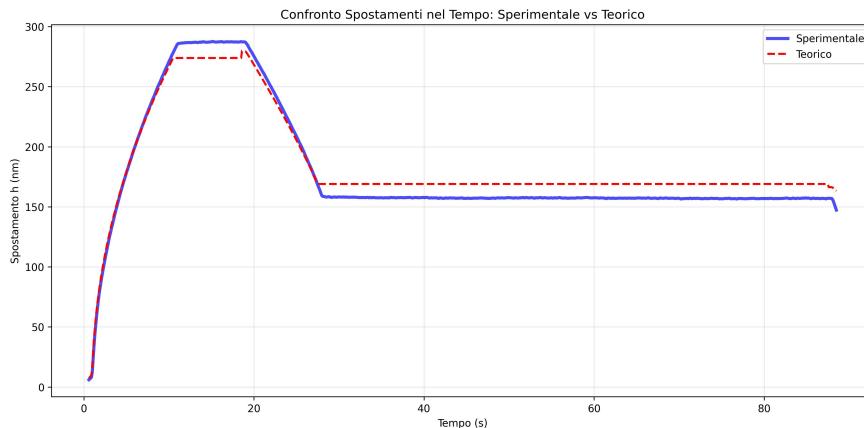


Figura 4.8. Profili di spostamento nel tempo sovrapposti.

Analizzando sia le figure che la tabella, otteniamo dei risultati ottimi, sia in quanto al confronto del calcolo degli spostamenti teorici, sia attraverso le metriche utilizzate. Infatti con un errore relativo medio del 6.28% e una correlazione del 0.9568, utilizzando gli intervalli del paragrafo precedente, possiamo dire con certezza che l'osservazione è in linea con il modello teorico.

Tabella 4.1. Riepilogo dati e metriche globali

<b>Informazioni Generali</b>	
Numero di punti analizzati	886
Range temporale	0.1 - 88.6 s
Range forza	0.0002 - 10.0074 N
Range spostamento	1.5 - 287.5 nm

<b>Metriche Globali</b>	
Errore relativo medio	6.28%
Correlazione forma	0.9568

Ovviamente in questo caso, essendo quest'ultima un'osservazione utilizzata per la stima del modello stesso, ci aspettiamo dei risultati molto elevati, che però possono molto diversi per nuove sperimentazioni.

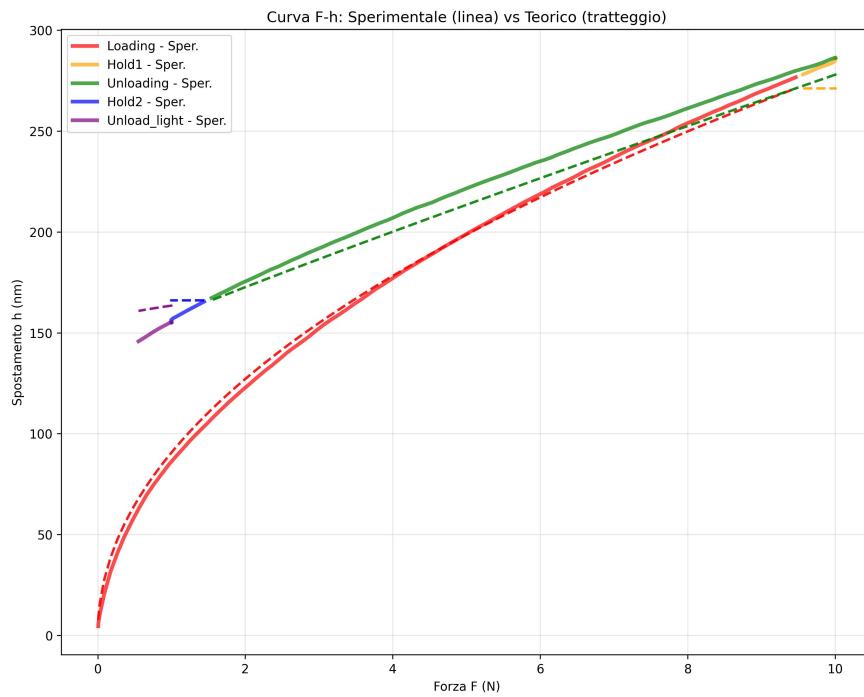


Figura 4.9. Profili di forza-spostamento sovrapposti.

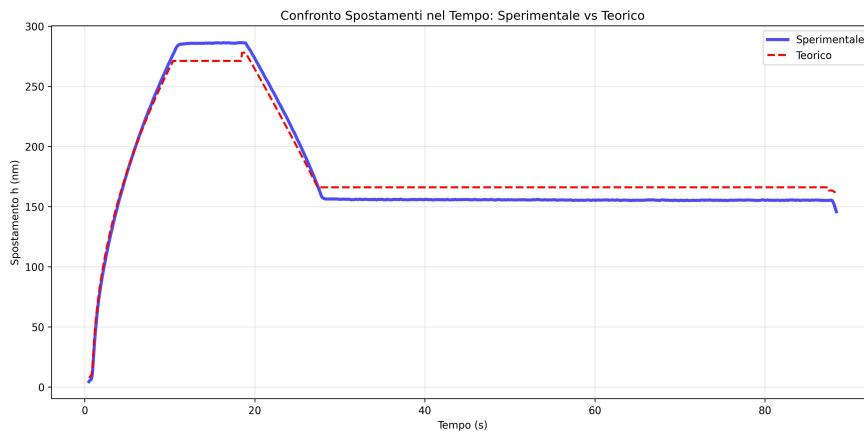


Figura 4.10. Profili di spostamento nel tempo sovrapposti.

#### 4.2.8 Dataset con controllo termico

Per il dataset con controllo termico è stato invece utilizzato il file di testo contenente la settima prova, di cui riportiamo i risultati, sia in figure che in tabella.

I risultati esposti sono molto simili al dataset precedente, sia per quanto riguarda i grafici sia per quanto riguarda le metriche, presentando nuovamente valori pari a 6.30% per l'errore relativo medio e 0.9527 per la

Tabella 4.2. Riepilogo dati e metriche globali

Informazioni Generali	
Numero di punti analizzati	884
Range temporale	0.1 - 88.4 s
Range forza	0.0007 - 10.0031 N
Range spostamento	1.4 - 286.4 nm
Metriche Globali	
Errore relativo medio	6.30%
Correlazione forma	0.9527

correlazione, il che come detto precedentemente sono ottimali. Nuovamente tutto il discorso dovrà essere ampliato quando il programma potrà essere testato con osservazioni nuove.

#### 4.2.9 Dataset teorico e sperimentale diversi

Come abbiamo sottolineato nelle due casistiche appena analizzate, i valori sia di Errore Relativo Medio, che di Correlazione, sono ottimali, avendo usato come esempi file che erano già stati usati per la stima dei parametri del modello teorico. Per questo motivo per verificare l'efficacia del codice creato per confrontare i nuovi dati sperimentali proponiamo un esempio nel quale abbiamo utilizzato il modello teorico con controllo termico, ma ipotizzando un nuovo dato proveniente da un'indentazione senza controllo termico.

Nello specifico verrà usato il file numero 37 dal dataset senza controllo termico, ipotizzando come detto un modello teorico inverso, presentiamo i risultati a seguire:

Si può chiaramente notare, sia dai grafici presenti nelle figure 4.11 e 4.12, sia nelle metriche globali della tabella sovrastante, che presentano differenze ben maggiori rispetto ai dati corretti. La metrica che è più evidente è l'Errore Relativo Medio, che è passato da valori del 5% e del 6% a un valore elevato del 9.81%. Per quanto riguarda la correlazione possiamo notare un leggero peggioramento, non superando la soglia di 0.95, come accaduto invece nei due casi precedenti. La correlazione rimane comunque elevata perché per quanto detto quando è stata descritta, essa fornisce un valore sull'andamento delle due funzioni in intervalli equivalenti, e entrambi i dataset in realtà seguono lo stesso andamento, a meno di davvero pochi punti e di conseguenza intervalli.

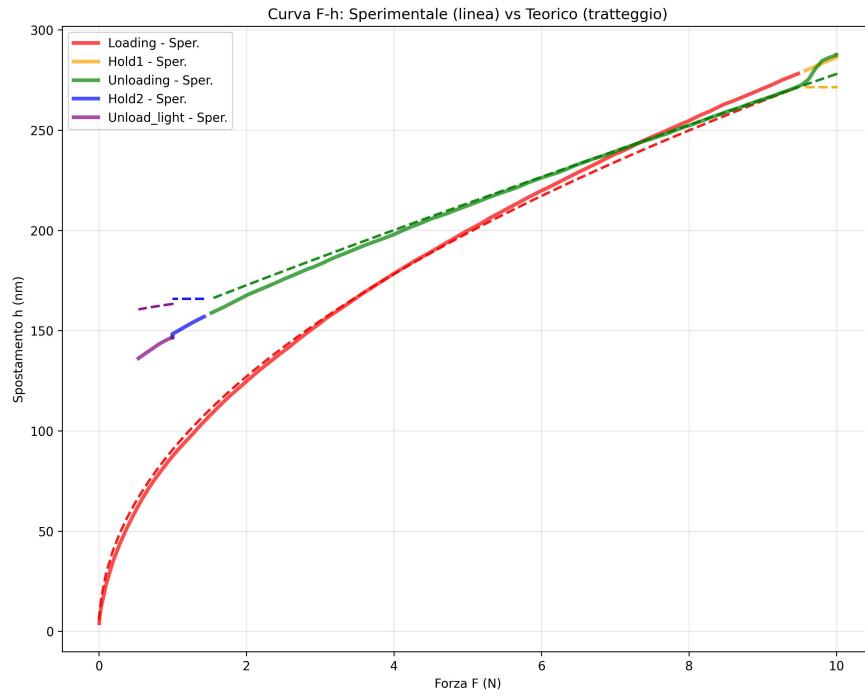


Figura 4.11. Profili di forza-spostamento sovrapposti.

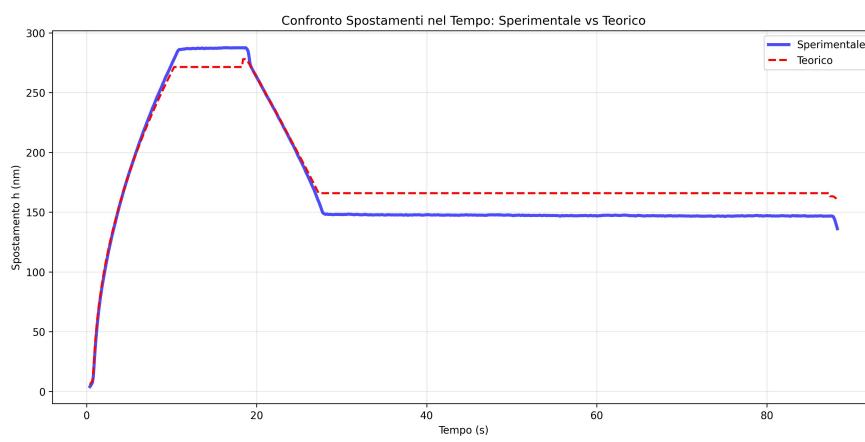


Figura 4.12. Profili di spostamento nel tempo sovrapposti.

Tabella 4.3. Riepilogo dati e metriche globali

<b>Informazioni Generali</b>	
Numero di punti analizzati	883
Range temporale	0.1 - 88.3 s
Range forza	0.0001 - 10.0008 N
Range spostamento	1.4 - 287.6 nm

<b>Metriche Globali</b>	
Errore relativo medio	9.81%
Correlazione forma	0.9451

# Bibliografia

- [1] James O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer, 1985.
- [2] Jose M. Bernardo and Adrian F. M. Smith. *Bayesian Theory*. Wiley, 2009.
- [3] Donald A. Berry. Bayesian clinical trials. *Nature Reviews Drug Discovery*, 5:27–36, 2006.
- [4] Michael Betancourt. A conceptual introduction to hamiltonian monte carlo. *arXiv preprint arXiv:1701.02434*, 2017.
- [5] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. A tutorial on bayesian model calibration. *Statistical Science*, 26(1):78–97, 2011.
- [6] Kenneth P Burnham and David R Anderson. Model selection and multimodel inference: a practical information-theoretic approach. *Springer*, 2002.
- [7] Bob Carpenter, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of Statistical Software*, 76:1–32, 2017.
- [8] George Casella and Roger L Berger. *Statistical inference*. Duxbury Press, Pacific Grove, CA, 2nd edition, 2002.
- [9] Jacob Cohen, Patricia Cohen, Stephen G West, and Leona S Aiken. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, 3rd edition, 2003.
- [10] Thomas S. Ferguson. A bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1(2):209–230, 1973.
- [11] International Organization for Standardization. Iso 23247-1:2021 automation systems and integration-digital twin framework for manufacturing. part 1: Overview and general principles, 2021.
- [12] Python Software Foundation. glob — unix style pathname pattern expansion, 2020. Python Standard Library Documentation.
- [13] Python Software Foundation. os — miscellaneous operating system interfaces, 2020. Python Standard Library Documentation.

- [14] Python Software Foundation. *pathlib — object-oriented filesystem paths*, 2020. Python Standard Library Documentation.
- [15] Python Software Foundation. *warnings — warning control*, 2020. Python Standard Library Documentation.
- [16] Alan E. Gelfand and Adrian F. M. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409, 1990.
- [17] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian Data Analysis*. CRC Press, 3rd edition, 2013.
- [18] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC, 3rd edition, 2014.
- [19] Andrew Gelman and Donald B Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472, 1992.
- [20] W. R. Gilks, Sylvia Richardson, and David Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC, 1996.
- [21] Michael Grieves. *Digital Twin: Manufacturing Excellence through Virtual Factory Replication*. Digital Manufacturing, 2014.
- [22] Charles R Harris, K Jarrod Millman, Stéfan J van der Walt, et al. Numpy: The fundamental package for scientific computing with python, 2020.
- [23] Frank H"artig, Karin Kniel, and Daniel Heißelmann. Das virtuelle koordinatenmessgerät—ein digitaler metrologischer zwilling. *Tm-Technisches Messen*, 90(9):548–556, 2023.
- [24] Edward P. Herbst. Bayesian analysis of garch and stochastic volatility models. *Journal of Financial Econometrics*, 2:173–205, 2000.
- [25] Matthew D Hoffman and Andrew Gelman. The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, 15(1):1593–1623, 2014.
- [26] John D Hunter. Matplotlib: A 2d graphics environment, 2007.
- [27] Harold Jeffreys. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A*, 186:453–461, 1946.
- [28] David Jones, Chris Snider, Aydin Nassehi, Jason Yon, and Ben Hicks. Characterising the digital twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, 29:36–52, 2020.
- [29] Michael I. Jordan, Zoubin Ghahramani, Tommi S. Jaakkola, and Lawrence K. Saul. An introduction to variational methods for graphical models. *Machine Learning*, 37:183–233, 1999.

- [30] Reji Jose and Seeram Ramakrishna. Materials 4.0: Materials big data enabled materials discovery. *Applied Materials Today*, 10:127–132, 2018.
- [31] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.
- [32] Pranav M Karve, Yongming Guo, Berkcan Kapusuzoglu, Sankaran Mahadevan, and Melanie A Haile. Digital twin approach for damage-tolerant mission planning under uncertainty. *Engineering Fracture Mechanics*, 225:106766, 2020.
- [33] James Kennedy and Russell Eberhart. Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks*, 1995.
- [34] Fook Fah Ling, Wen-An Lai, and David A Lucca. *Fundamentals of Surface Mechanics*. Springer Science and Business Media, 2012.
- [35] David A Lucca, Konrad Herrmann, and Michael J Klopfstein. Nanoindentation: Measuring methods and applications. *CIRP Annals*, 59(2):803–819, 2010.
- [36] Giacomo Maculotti, Gianfranco Genta, and Maurizio Galetto. Single-step calibration method for nano indentation testing machines. *CIRP Annals*, 69:429–432, 2020.
- [37] Giacomo Maculotti, Gianfranco Genta, and Maurizio Galetto. An uncertainty-based quality evaluation tool for nanoindentation systems. *Measurement*, 225:113974, 2024.
- [38] Maria Chiara Magnanini and Tullio AM Tolio. A model-based digital twin to support responsive manufacturing systems. *CIRP Annals*, 70(1):353–356, 2021.
- [39] Wes McKinney et al. pandas: powerful python data analysis toolkit, 2010.
- [40] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [41] Douglas C Montgomery. *Introduction to Statistical Quality Control*. John Wiley & Sons, 8th edition, 2017.
- [42] Radford M. Neal. Bayesian learning for neural networks. *PhD thesis, University of Toronto*, 1996.
- [43] Warren C Oliver and George M Pharr. An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. *Journal of materials research*, 7(6):1564–1583, 1992.

- [44] Warren C Oliver and George M Pharr. Measurement of hardness and elastic modulus by instrumented indentation: Advances in understanding and refinements to methodology. *Journal of Materials Research*, 19(1):3–20, 2004.
- [45] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, et al. Scikit-learn: Machine learning in python, 2011.
- [46] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.
- [47] Industry Research. Digital twin technology benefits and applications. Web search results, 2024.
- [48] Gideon Schwarz. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.
- [49] David J Spiegelhalter, Nicola G Best, Bradley P Carlin, and Angelika Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the royal statistical society: Series b (statistical methodology)*, 64(4):583–639, 2002.
- [50] Rainer Storn and Kenneth Price. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341–359, 1997.
- [51] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 27(5):1413–1432, 2017.
- [52] Aki Vehtari, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul-Christian Bürkner. Rank-normalization, folding, and localization: An improved  $\hat{R}$  for assessing convergence of mcmc. *Bayesian Analysis*, 16(2):667–718, 2021.
- [53] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, et al. Scipy: Open source scientific tools for python, 2020.
- [54] Sumio Watanabe. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11:3571–3594, 2010.
- [55] Cort J Willmott and Kenji Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate Research*, 30(1):79–82, 2005.
- [56] Sungmin Yoon and Jihun Koo. In situ model fusion for building digital twinning. *Building and Environment*, 243:110652, 2023.
- [57] Yixuan Zheng, Shijie Wang, Qican Li, and Benyong Li. Fringe projection profilometry by conducting deep learning from its digital twin. *Optics Express*, 28(24):36568, 2020.