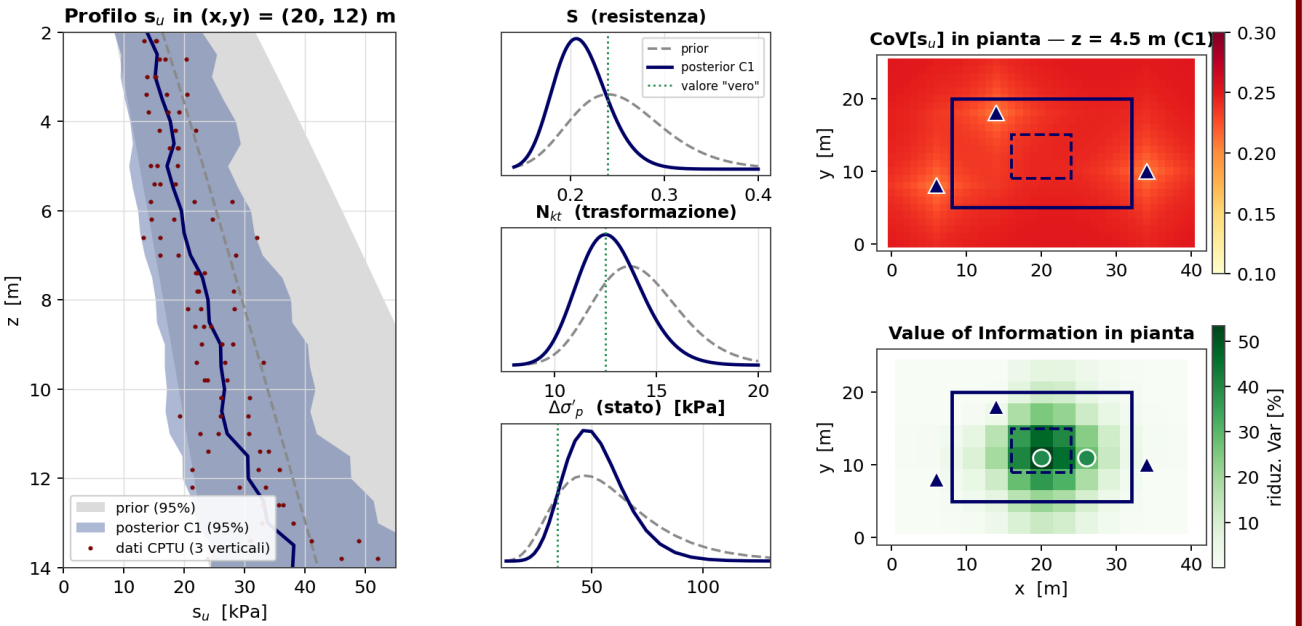
UNIVERSITÀ DEGLI STUDI FIRENZE
DICEA
Dipartimento di Ingegneria Civile e Ambientale – Sezione Geotecnica
Department of Civil and Environmental Engineering – Geotechnical Group



estratto dal codice del caso di studio (Python + numpy/scipy)

```
import numpy as np;  from scipy.stats import norm, multivariate_normal as mvn

# 1 - dati della campagna e modello fisico (SHANSEP)
P, y, H, R = leggi_campagna(cptu, laboratorio)      # 154 misure
sv0   = 10.0 + 8.2 * P[:, 2]                        # tensione efficace
g_ocr = lambda dsp: m * np.log((sv0 + dsp) / sv0)  # termine di stato

# 2 - struttura di correlazione: kriging bayesiano sul residuo
def acf(A, B):
    dxy = dist_pianta(A, B);  dz = dist_vert(A, B)
    return np.exp(-2 * dxy / theta_h - 2 * dz / theta_v)
C  = sig_w**2 * acf(P, P) + R                       # + errore di misura
Ci = np.linalg.inv(C);  M = H @ Sig0 @ H.T + C

# 3 - posterior di Theta = {S, N_kt, dsigma_p} per marginalizzazione
Sg = np.linalg.inv(iSig0 + H.T @ Ci @ H)
for k, dsp in enumerate(griglia_dsp):
    gk        = g_ocr(dsp)
    logev[k] = mvn.logpdf(y, mean=H @ mu0 + gk, cov=M)
    mus[k]   = Sg @ (iSig0 @ mu0 + H.T @ Ci @ (y - gk))
logev += norm.logpdf(np.log(sigma_p_oed), np.log(griglia_dsp), s_oed)
w   = np.exp(logev + log_prior);  w /= w.sum()      # pesi a posteriori
mu  = (w[:, None] * mus).sum(0)                     # E[ln S], E[ln N_kt]

# 4 - campo condizionato (kriging) e mappa del CoV
K   = sig_w**2 * acf(Q, P)                          # Q: nodi di griglia
mw  = K @ Ci @ (y - (H @ mu + g_ocr(dsp_hat)))      # media di kriging
s2  = sig_w**2 - np.einsum('ij,jk,ik->i', K, Ci, K) + Sig[0,0] + var_t
su  = np.exp(np.log(sv0_Q) + t_bar + mu[0] + mw + s2 / 2)
CoV = np.sqrt(np.exp(s2) - 1)                       # criterio di arresto

# 5 - propagazione ai valori rappresentativi (EC7-2G)
su_vol = monte_carlo(mw, Vw, mu, w, Q_uls, n=3000) # medie sul volume
X_rep  = np.quantile(su_vol, 0.05)                  # Caso A / Caso B
```

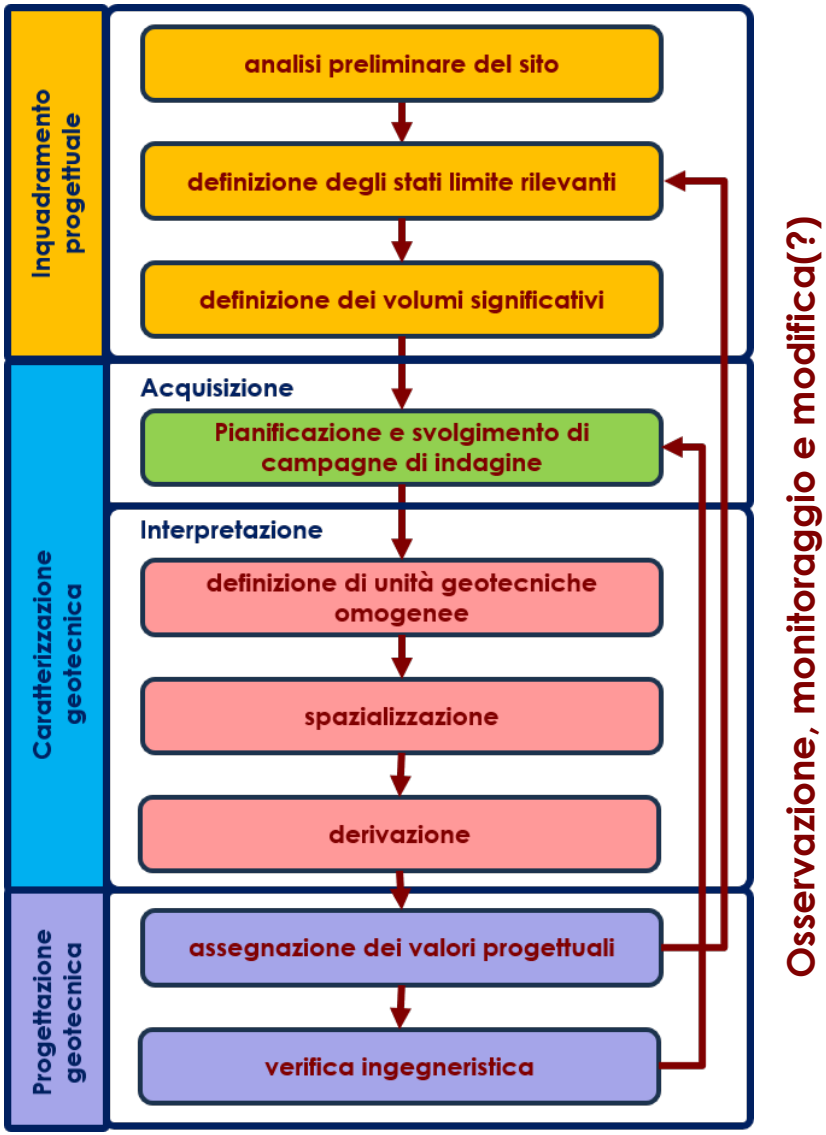
Caratterizzazione geotecnica adattativa mediante machine learning “physics-informed” e approccio Bayesiano: princìpi e applicazione

Marco UZIELLI

Associazione Geotecnica Italiana – Giornata di studio “Intelligenza artificiale e ingegneria geotecnica”
Roma, 21 settembre 2026

La buona pratica geotecnica: il processo e gli ostacoli

- Il processo della buona pratica geotecnica è iterativo
- Ogni fase e ogni operazione introduce incertezze specifiche, intrinseche al processo conoscitivo
- Le incertezze si accumulano progressivamente, fino ai valori progettuali
- La quantificazione delle incertezze è di fatto una "richiesta" normativa (assegnazione dei valori caratteristici)
- Gli scenari "tipici" di caratterizzazione geotecnica (dati limitati, spazialmente distanti, puntuali o al massimo 1D) non consentono la definizione affidabile di valori statisticamente derivati all'interno di un "volume significativo" come da Normativa
- La prossima evoluzione normativa (EC7-2G) assegna un ruolo centrale alla definizione delle incertezze nel modello geotecnico e apre ad analisi pienamente probabilistiche (Reliability-based design - RBD)



Le incertezze geotecniche: glossario e classificazione

Incertezza aleatoria, dovuta alla variabilità "reale" (spaziale, temporale) delle proprietà compositive e meccaniche dei terreni

Incertezza epistemica, dovuta all'imperfezione dei processi di acquisizione, alla quantità finita di dati acquisibili, al livello di approssimazione modellistica, al passaggio da parametri di ingresso misurati a parametri derivati e al confronto tra i risultati dei modelli e i parametri in uscita misurati localmente

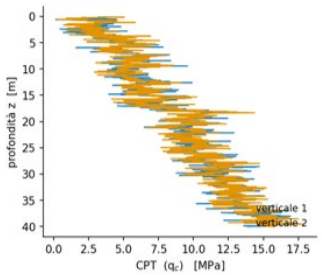
variabilità intrinseca (spaziale, temporale)

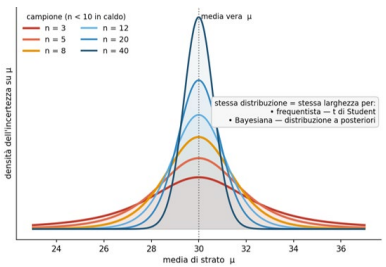
incertezza di misura

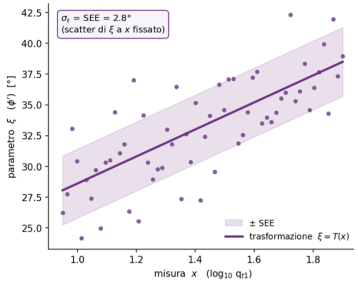
incertezza statistica

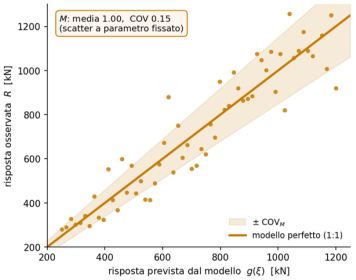
incertezza di trasformazione

incertezza di modello









"variabilità reale"

"incertezza di indagine"

"incertezza di derivazione"

REALTÀ FISICA preesiste al processo	**campagna di indagine**	**spazializzazione** → **derivazione**
	ACQUISIZIONE	**INTERPRETAZIONE**

Modellare le incertezze: la via frequentista e le sue limitazioni

Gran parte della pratica normata adotta, più o meno esplicitamente, l'approccio frequentista, la cui applicazione rigorosa richiede un numero di dati sufficiente e, spesso, l'ipotesi di una distribuzione (tipicamente gaussiana).

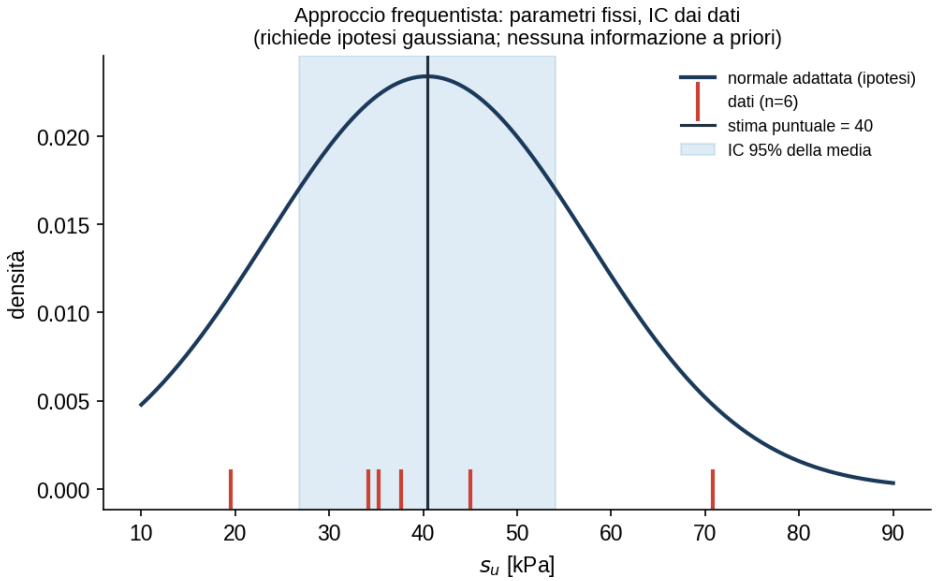
Nella pratica i dati per strato sono pochi e le distribuzioni dei parametri sono spesso non gaussiane (asimmetriche, limitate).

Con pochi dati, la via frequentista produce un valore puntuale e un intervallo di confidenza, assumendo una forma gaussiana e senza informazione a priori.

- L'intervallo di confidenza dipende dalla numerosità: con pochi dati è ampio.
- La validità poggia sull'ipotesi gaussiana, spesso non verificata.

Queste incongruenze sono convenzionalmente accettate: pertanto, si applicano spesso procedure le cui ipotesi non sono verificate!

Le stime risultano instabili e gli intervalli ampi, ma ciò resta non dichiarato.



Modellare le incertezze: l'alternativa Bayesiana

L'approccio Bayesiano supera molte delle limitazioni di quello frequentista descrivendo i parametri come distribuzioni e aggiornando esplicitamente conoscenze o ipotesi preliminari con l'evidenza.

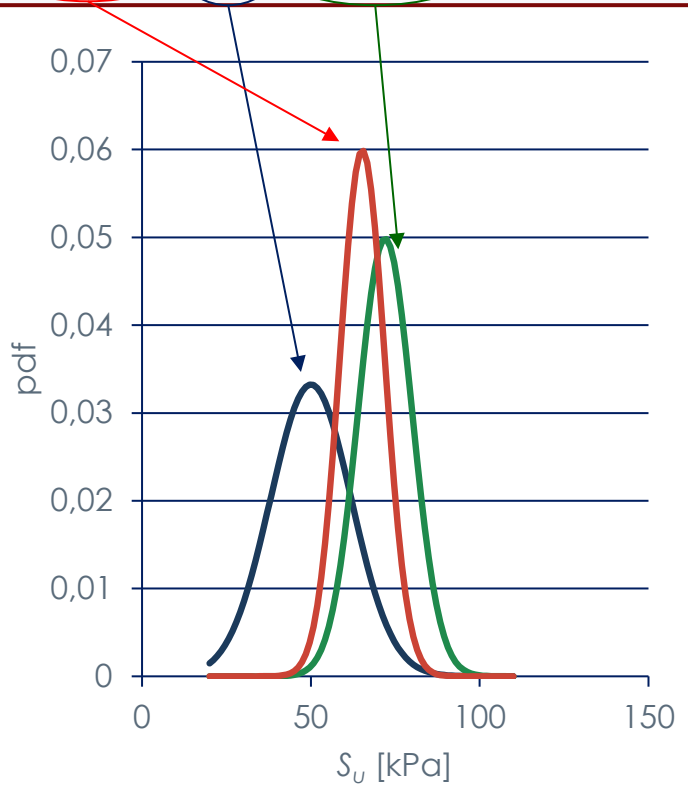
- I parametri non sono valori fissi ignoti, bensì variabili aleatorie descritte da distribuzioni di probabilità.
- Si parte da una conoscenza a priori e la si aggiorna con i dati osservati (likelihood) per ottenere una distribuzione a posteriori.
- L'approccio Bayesiano non impone sempre una specifica distribuzione ai dati; si possono adottare distribuzioni diverse e coerenti con ciascuno dei parametri oggetto di indagine (limitate, asimmetriche, ecc.).

Teorema di Bayes

Rev. Thomas Bayes (1702–1761)

$$P(H|D) = \frac{P(H) \cdot P(D|H)}{P(D)}$$

posterior = *prior* × *likelihood* / *evidenza*

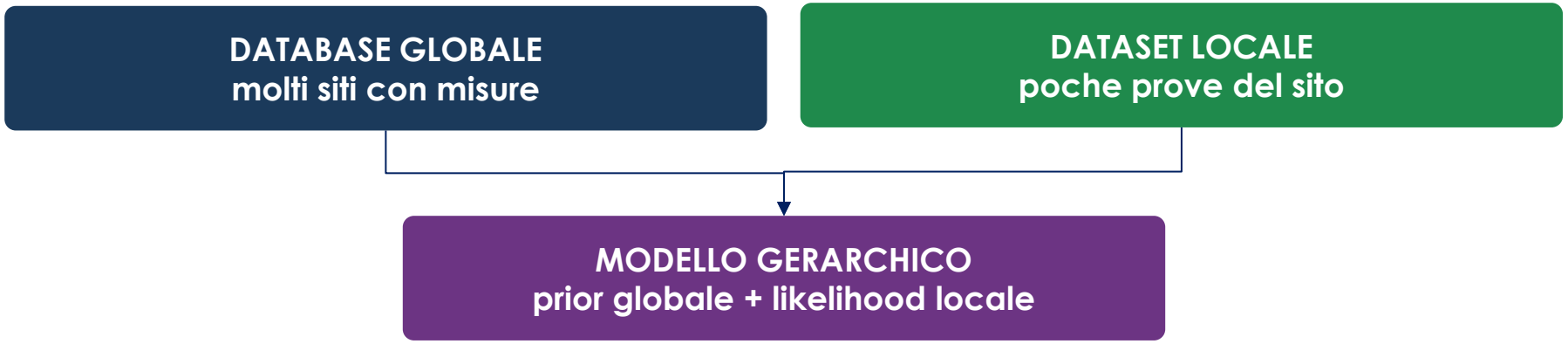


Modellare le incertezze: l'alternativa Bayesiana

Hierarchical Bayesian Modeling (HBM)

La modellazione Bayesiana gerarchica (HBM) mette in sinergia la conoscenza generale e i dati del singolo sito.

- Database globale: conoscenza accumulata su molti siti analoghi (database gratuiti).
- Dataset locale: i dati del sito in esame.
- L'HBM aggiorna una prior globale con una likelihood locale: i siti oggetto di indagine (tipicamente) poveri di dati si appoggiano in maniera ragionata a dataset "globali", pesandoli adeguatamente in relazione alla loro affinità e riducendo quindi l'incertezza epistemica statistica



I database multivariati globali (es. ISSMGE TC304) sono la base empirica dei modelli di trasformazione e delle distribuzioni a priori: quantificano correlazioni e dispersione fra grandezze misurate e parametri derivati, e consentono stime stabili dove i dati sito-specifici sono scarsi. Sono risorse aperte, alle quali è importante contribuire.

Modellare le incertezze: l'alternativa Bayesiana

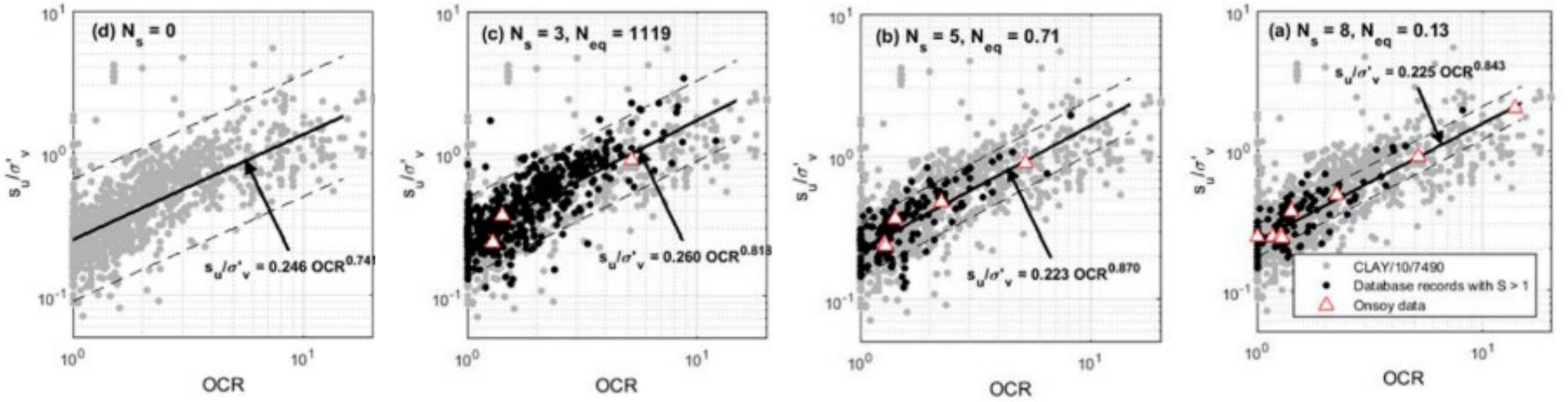
Hierarchical Bayesian Modeling (HBM)

L'approccio HBM è molto utile in quanto si adatta bene ai dati geotecnici che sono multivariati, incerti e unici, sparsi, incompleti, corrotti, spazialmente variabili, tridimensionali (acronimo MUSIC-3X).

Numerosi contributi di letteratura mostrano che la corretta trattazione delle incertezze (es. utilizzo dell'approccio Bayesiano in luogo di quello frequentista) consente di utilizzare in maniera efficiente dati sito-specifici per definire il carattere "locale" di modelli di trasformazione e ridurre la relativa incertezza.

Esempio: aggiornamento sito-specifico di correlazione "globale" tra rapporto di sovraconsolidazione OCR e coesione non drenata s_u, con riduzione dell'incertezza di trasformazione (Ching & Phoon 2019)

Solo dataset "globale" → **3 dati locali** → **6 dati locali** → **9 dati locali**



Ching & Phoon (2019)

Motivazione e tesi della presentazione

Rilevanza dell'AI e del ML in geotecnica nello scenario attuale e futuro

- incremento della capacità di acquisizione e interpretazione di dati geotecnici e di implementazione di metodi Bayesiani computazionali (fattore tecnologico)
- necessità di modellare le incertezze (fattore normativo)
- crescente applicazione di modelli data-driven per la stima di parametri geotecnici da dati sperimentali

I limiti del ML "convenzionale"

- difficile interpretazione e trasferibilità nella pratica geotecnica professionale
- le leggi fisiche non sono incorporate
- l'incertezza non è necessariamente quantificata rigorosamente

La proposta

- approccio integrato **physics-informed + Bayesiano** nel quale il ML supporta un processo di caratterizzazione sequenziale e adattativo, guidato dal professionista e dalla sua conoscenza geotecnica

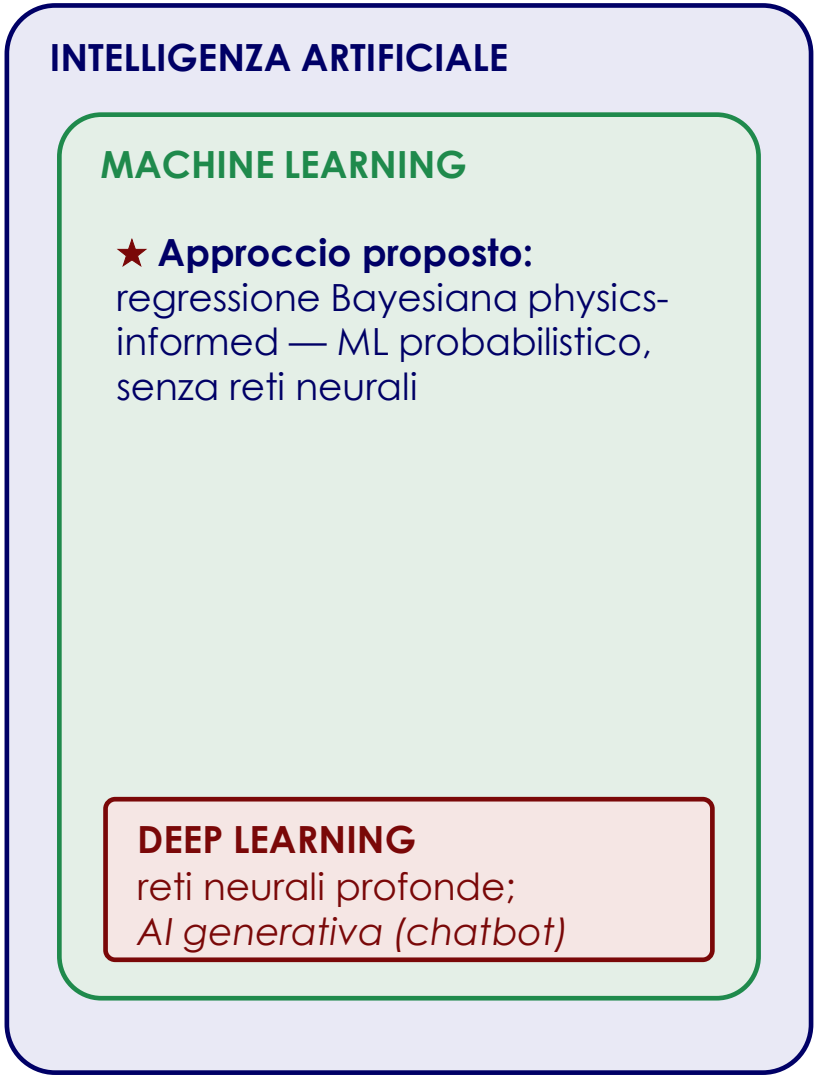
La tesi di fondo

L'AI non sostituisce l'ingegnere geotecnico: lo supporta attraverso un approccio efficiente all'interpretazione dei dati, coerente con i princìpi della geotecnica e mirato alla quantificazione esplicita dell'incertezza.

Intelligenza artificiale e machine learning: definizioni

- **Intelligenza artificiale (AI)**: la disciplina, ampia — sistemi che svolgono compiti che richiederebbero intelligenza umana
- **Machine learning (ML)**: un sottoinsieme dell'AI — algoritmi che migliorano le proprie prestazioni imparando dai dati anziché seguire solo regole pre-programmate (es. calibrazione di modelli e stima di funzioni e parametri a partire da esempi)
- **Deep learning**: un sottoinsieme del ML — reti neurali profonde (milioni di parametri, necessità di dataset di apprendimento molto grandi)
- **AI generativa (chatbot)**: applicazione del deep learning che produce testi, immagini, video, ecc.

L'approccio proposto ricade nel ML: apprende dai dati, ma con modelli probabilistici a pochi parametri e interpretabili dal punto di vista ingegneristico (non con reti neurali)



Limiti del machine learning puramente data-driven

Capacità del ML data-driven

- Apprende relazioni tra grandezze direttamente dai dati, anche complesse (non lineari e multivariate); elabora rapidamente grandi volumi; interpola con flessibilità
- Ma i dati geotecnici sono "**MUSIC-3X**": **M**ultivariati, incerti e unici (**U**ncertain and unique), **S**parsi, **I**ncompleti, **C**orrotti, spazialmente variabili, **3D** → il contesto peggiore per il ML "convenzionale"

Limiti strutturali del ML convenzionale in ambito progettuale

- **1. Opacità (black-box):** difficile spiegare perché il modello fornisce una certa stima → incompatibile con un processo che deve essere tracciabile, motivato e documentato (EC7-2G)
- **2. Scarsa trasferibilità:** un modello addestrato su altri siti incorpora la sito-specificità di quei siti; applicarlo acriticamente equivale a usare una correlazione fuori dal suo dominio di validità
- **3. Nessuna quantificazione rigorosa dell'incertezza:** una stima puntuale senza distribuzione non consente l'applicazione rigorosa dell'approccio semiprobabilistico (né, a maggior ragione, di quello probabilistico)

Il ML "convenzionale" equivale a una persona abilissima nel calcolo matematico, ma che non conosce la geotecnica e non dichiara quanto è sicura delle proprie stime!

Machine learning "physics-informed": definizione e motivazioni

Principio

- I modelli di apprendimento non sono liberi di formulare relazioni arbitrarie: sono vincolati o guidati dalla conoscenza delle leggi fisiche che governano il comportamento dei terreni — vincoli di ammissibilità dei parametri, coerenza con relazioni costitutive e modelli di trasformazione, struttura del modello meccanicamente sensata

Vantaggi per la pratica geotecnica

- Estrapolazione più sicura perché i dati geotecnici sono "MUSIC-3X" e sito-specifici: la fisica evita di andare alla deriva fuori dal dominio dei dati osservati
- Interpretabilità e tracciabilità: risultati leggibili nel linguaggio dell'ingegneria geotecnica
- Coerenza fisica garantita: nessuna stima fisicamente "impossibile" (resistenze negative, moduli non plausibili)
- Efficienza con pochi dati: un modello che "parte già sapendo la fisica" richiede meno dati di addestramento (fondamentale con i dataset geotecnici tipici)
- Sinergia naturale con l'approccio Bayesiano: la fisica struttura il modello, l'approccio Bayesiano ne quantifica e aggiorna l'incertezza

Nella modalità "physics-informed", il ML smette di essere una black-box e diventa uno strumento ingegneristico trasparente, coerente e difendibile.

L'algoritmo adattativo: il ciclo operativo

Step 0: PRIOR — distribuzioni iniziali dei parametri incogniti, da dataset globali (HBM) e giudizio esperto

↓

Step 1: ACQUISIZIONE — prove in sito, sondaggi con prelievo di campioni, prove di laboratorio; in esercizio: monitoraggio

↓

Step 2: LIKELIHOOD physics-informed — trend fisico e modelli di trasformazione + spazializzazione (kriging Bayesiano)

↓

Step 3: POSTERIOR — aggiornamento Bayesiano iterativo: la posterior di un ciclo è la prior del ciclo successivo

↓

Step 4: PROPAGAZIONE — Monte Carlo sui modelli di stato limite: valori caratteristici (Casi A/B)

↓

incertezza ≤ soglia di tollerabilità per ogni stato limite?

NO → Step 5: PIANIFICAZIONE (VoI) — ubicazione e acquisizione supplementare dove la riduzione attesa di varianza per unità di costo è massima

SÌ → STOP — valori progettuali e verifica (RBD / LRFD); in esercizio: monitoraggio come nuova likelihood

L'ingegnere geotecnico valida e decide a ogni passo: unità omogenee, prior, campagna, valori finali.

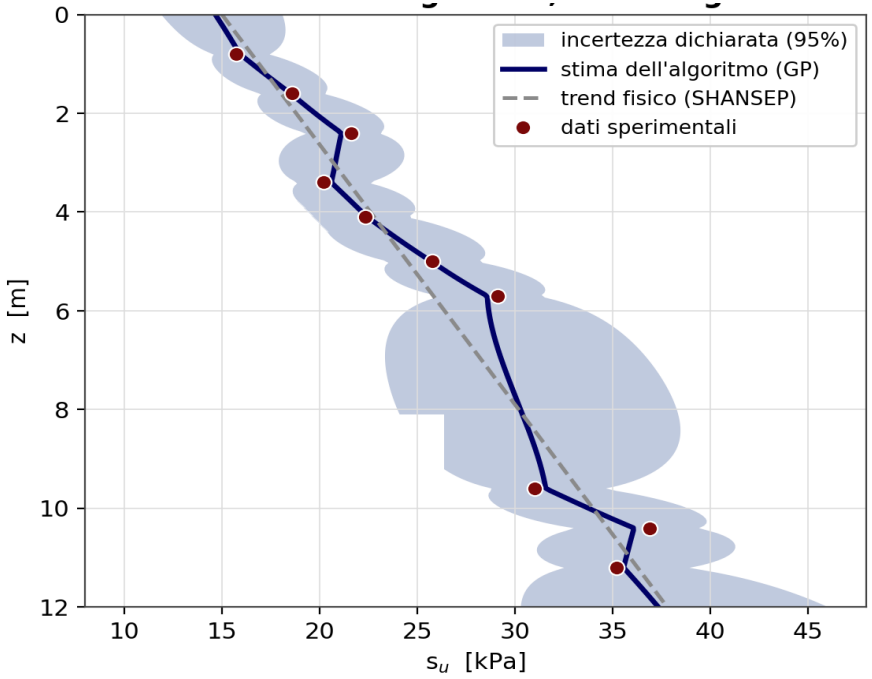
Che cos'è, in concreto

- Non un pacchetto software commerciale, ma un metodo di calcolo implementato in uno script ispezionabile e modificabile
- Ingressi: fisica e conoscenza geotecnica, dataset globali (prior), dati locali. Uscita: distribuzioni aggiornate
- "Apprendere dai dati" = stimare parametri e incertezze; la fisica vincola la forma del modello

L'algoritmo: regressione Bayesiana "istruita" dalla fisica

- **Algoritmo composito**: il nucleo è la regressione Bayesiana con processi gaussiani (GP), generalizzazione del kriging, in uso in geotecnica da decenni
- **Il processo gaussiano stima la forma del profilo dai dati**; la fisica (modello SHANSEP, domini di ammissibilità) vincola le stime ad andamenti "geotecnicamente plausibili"
- **L'"addestramento" non è una black-box**: coincide con il calcolo della posterior — in forma chiusa in log-spazio o via Monte Carlo — integralmente riproducibile e fondato su dati geotecnici "veri", ancorché non locali
- **Con i piccoli dataset locali del singolo sito (MUSIC-3X)** i modelli probabilistici trasparenti sono preferibili al deep learning (meno parametri, più interpretabilità; i dataset globali multi-sito entrano come prior)

Il risultato è una stima corredata della propria incertezza, direttamente utilizzabile nel quadro attuale (NTC2018) e futuro (EC7-2G).



Presso i punti misurati la banda di incertezza è stretta perché l'evidenza aiuta; lontano dai dati essa si allarga, ma il vincolo fisico mantiene la stima su un andamento meccanicamente ammissibile.

Pianificazione della campagna di acquisizione

La campagna di indagine si pianifica in funzione del progetto!

- Caratteristiche dell'opera (importanza, geometria, carichi).
- Gli stati limite ritenuti rilevanti in fase di inquadramento progettuale (es. i cedimenti richiedono verticali profonde; lo scorrimento richiede verticali ravvicinate). Tale scelta può variare nel corso del processo geotecnico e la campagna di acquisizione può essere integrata da nuove indagini!
- Il modello geologico: stratigrafia attesa, direzioni di variabilità, elementi rilevanti da considerare preliminarmente alla campagna.
- Le incertezze presumibili che si aggiungeranno a quelle dei parametri nella fase di interpretazione dei dati (le incertezze si accumulano progressivamente!)

Numero e tipologia delle prove: riformulazione del quesito

In assoluto, una formula univocamente corretta per il tipo e il numero "giusto" di prove non esiste: in geotecnica nessuna regola tabellare ha fondamento generale.

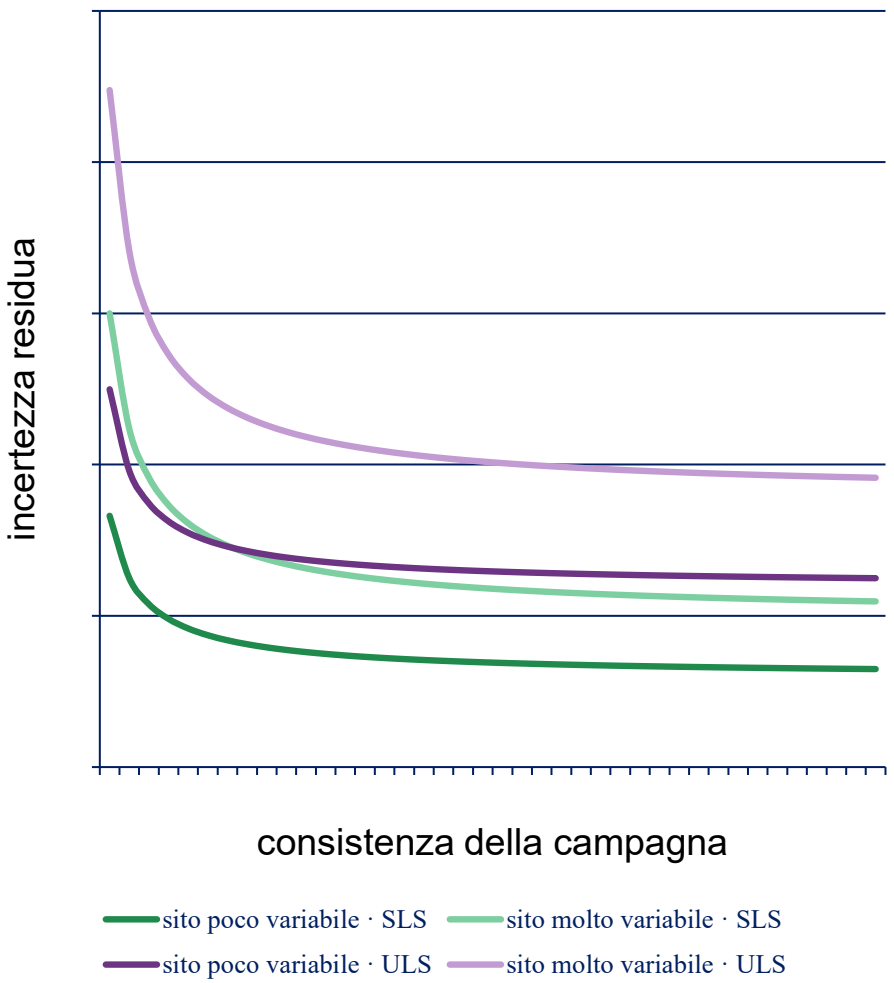
L'obiettivo è quello di portare l'incertezza progettuale entro limiti di accettabilità per tutti gli stati limite di interesse!

Il numero "ottimale" dipende:

- dalla variabilità spaziale delle proprietà geotecniche del sito: più il terreno varia, più informazione serve.
- dallo stato limite: la campagna si calibra sul meccanismo che genera maggiore incertezza.

"Quante prove, dove, a quale profondità": la risposta sta nella domanda progettuale, non (solo) nel terreno.

Principio di non saturazione: le incertezze si accumulano lungo la catena; non è opportuno partire già al limite della tollerabilità.



Caso di studio — Il sito e il problema progettuale

Caso realistico, sviluppato con finalità dimostrative.

L'opera e il sito

- Fondazioni superficiali di un edificio su deposito alluvionale eterogeneo: alternanze sabbioso-limose e argillose, variabilità spaziale marcata in verticale, falda prossima al piano di posa

Stati limite dichiarati ex ante (richiesti dal quadro vigente; ruolo centrale nell'EC7-2G)

- **ULS – capacità portante:** meccanismo tendenzialmente "locale" → alta sensibilità alla variabilità spaziale (EC7-2G: Caso B)
- **SLS – cedimenti:** meccanismo "non locale", mediato su un volume esteso → bassa sensibilità (EC7-2G: Caso A)

Conseguenza (la "dinastia modellistica")

- Un unico modello del sottosuolo, ma **due modelli geotecnici progettuali distinti**: volume significativo, riduzione della varianza e valori caratteristici diversi per ciascuno stato limite

Chiedere "il" parametro di progetto del sito è una domanda mal posta: i valori progettuali dipendono dallo stato limite.

Caso di studio — I modelli degli stati limite

Il modello di verifica determina la funzione dei parametri richiesta, il volume di riferimento e la statistica rappresentativa.

ULS — capacità portante (meccanismo "locale")

- **Modello (non drenato):**

$$q_{\lim} = N_c \cdot s_u + q_0,\ N_c = \pi + 2$$

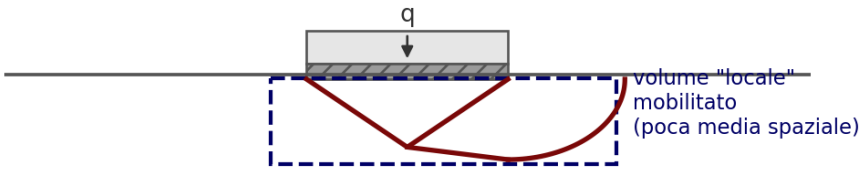
- **Grandezza richiesta:** su sul volume locale mobilitato dal cinematismo, sotto il piano di posa
- **Ruolo di Θ:** S, Δσ'p e Nkt determinano livello, forma e interpretazione del profilo; il piccolo volume mobilitato media poco la variabilità → varianza elevata
- **Valore caratteristico (EC7-2G: Caso B):** frattile 5% della distribuzione locale

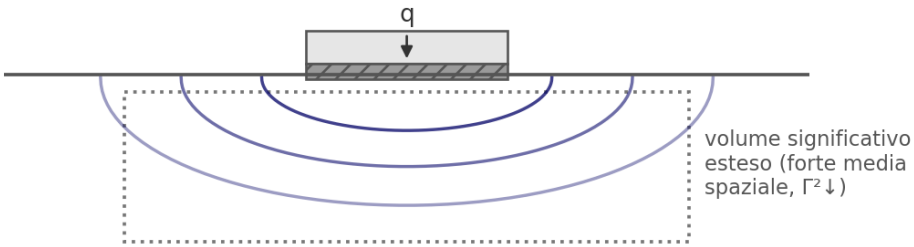
SLS — cedimenti (meccanismo "non locale")

- **Modello (elastico):**

$s = \frac{q \cdot B \cdot (1-\nu^2) \cdot I_s}{E_u}$, con $E_u = K_E \cdot s_u$ (KE assunto)

- **Grandezza richiesta:** media di su sul volume significativo esteso (bulbo tensionale)
- **Ruolo di Θ:** stessi parametri — e Δσ'p governa anche la rigidezza tramite OCR; il volume esteso media la variabilità spaziale → varianza ridotta ($\Gamma^2\downarrow$)
- **Valore caratteristico (EC7-2G: Caso A):** stima cauta (frattile 5%) della media spaziale





Un'unica caratterizzazione multivariata alimenta entrambe le verifiche, con valori caratteristici dati da statistiche differenziate in ragione dei meccanismi.

Caso di studio — Struttura del modello geotecnico

- **Struttura gerarchica in log-spazio:**

$$\ln s_{\mathrm{u,OC}} = \ln[S \cdot \sigma'_{\mathrm{v0}} \cdot OCR^m] + w(x,z)$$

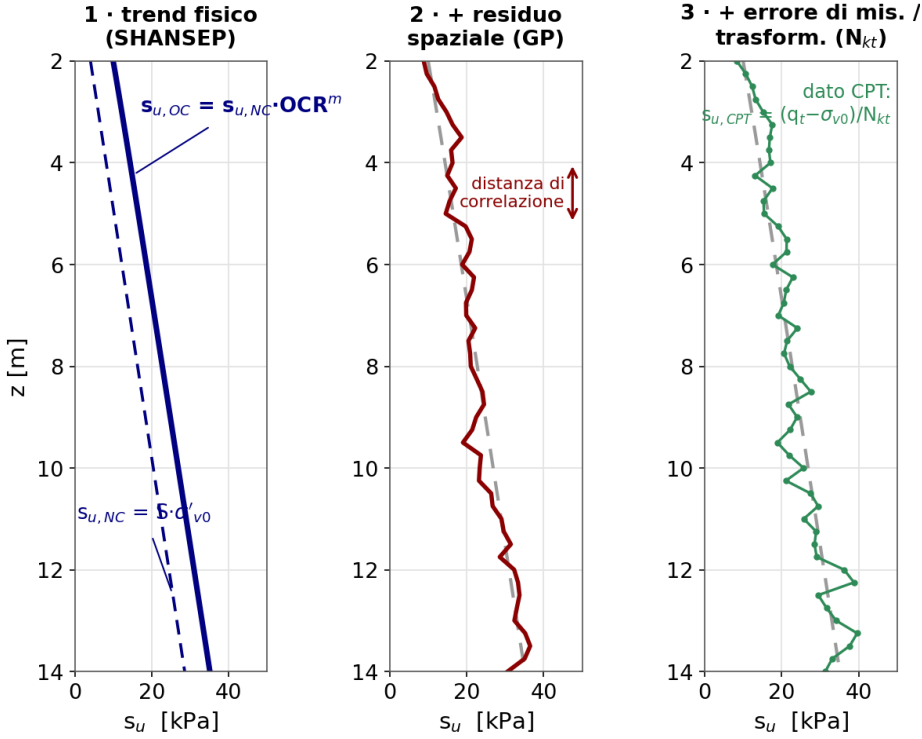
(trend fisico SHANSEP: $s_{\mathrm{u,NC}} = S \cdot \sigma'_{\mathrm{v0}}$; $s_{\mathrm{u,OC}} = s_{\mathrm{u,NC}} \cdot OCR^m$; w = residuo spaziale)

- **Il residuo w è un processo gaussiano a media nulla: correlazione** $\rho(\Delta x, \Delta z)$ con distanze di correlazione verticale e orizzontale (variabilità aleatoria); il suo condizionamento alle misure è la spazializzazione (kriging)
- **La misura locale entra nella likelihood:**

$$s_{\mathrm{u,CPT}} = \frac{q_{\mathrm{t}} - \sigma_{\mathrm{v0}}}{N_{\mathrm{kt}}}$$

con incertezza di trasformazione (su N_{kt}) e di misura (ε) esplicite

- **Le incognite** $\Theta = \{S, N_{\mathrm{kt}}, \Delta\sigma'_{\mathrm{p}}, \ldots\}$: tutte distribuzioni che si aggiornano iterativamente dai dati, non valori singoli né valori di letteratura non sito-specifici!
- **La fisica vincola il trend e i domini ammissibili:** l'apprendimento resta compatibile con la meccanica dei terreni



Ogni componente dell'incertezza ha collocazione esplicita nel modello, senza alcuna "black-box" geotecnica!

Caso di studio — Step 0: modello concettuale e prior

Modello geologico concettuale

- Stratigrafia attesa, direzioni di variabilità, elementi da ipotizzare prima dell'indagine (lezione di Terzaghi: partire sempre dalla geologia)

Assegnazione delle distribuzioni a priori (prior)

- Ai principali parametri geotecnici, combinando: **conoscenza geologica del sito**, **esperienza del professionista**, **letteratura**, **database globali** (prior gerarchica HBM da molti siti analoghi)
- Distribuzioni coerenti con la natura fisica dei parametri: limitate, eventualmente asimmetriche — nessuna gaussiana imposta per comodità

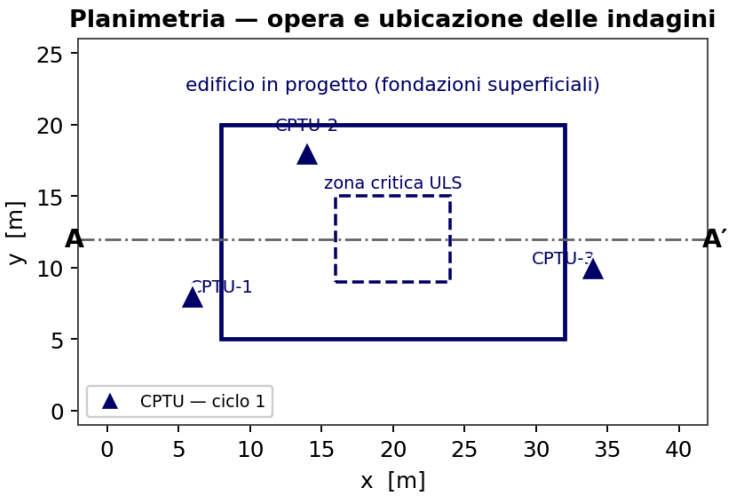
Il giudizio ingegneristico entra qui — formalmente

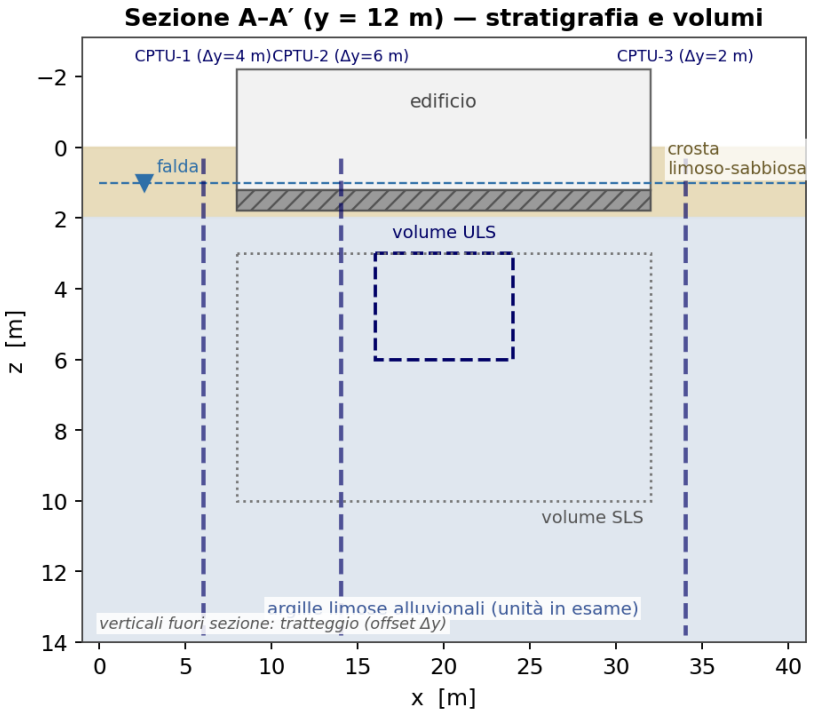
- La prior è la sede **dichiarata e tracciabile** del giudizio esperto: non più assegnazione soggettiva mascherata, ma ipotesi esplicita che i dati potranno correggere

Stato dell'incertezza allo Step 0: elevata, ma dichiarata e quantificata. Su questa base — modello geologico, stati limite, opera — si pianifica la campagna iniziale.

Caso di studio — Planimetria e campagna iniziale

La campagna iniziale discende dal modello geologico concettuale, dagli stati limite dichiarati e dalla geometria dell'opera, prima di ogni misura: la zona critica per l'ULS è individuata ex ante. Sezione A–A′ di riferimento a y = 12 m.



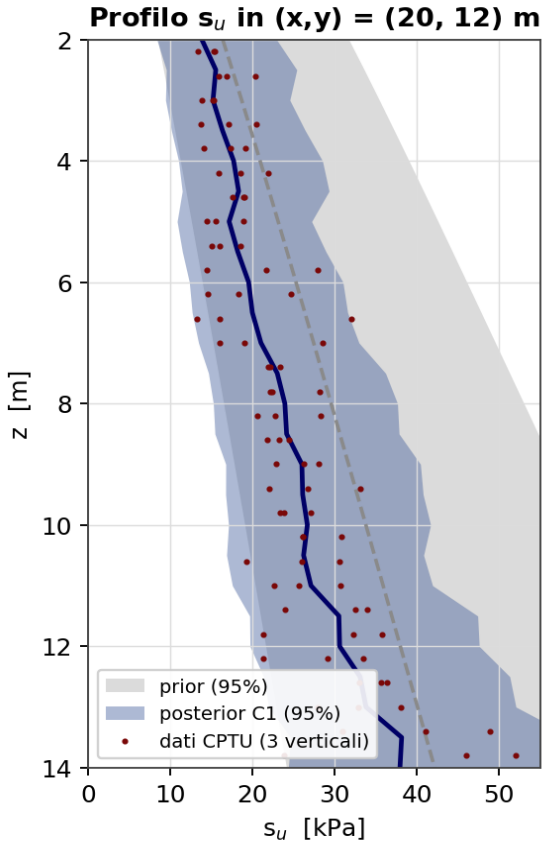


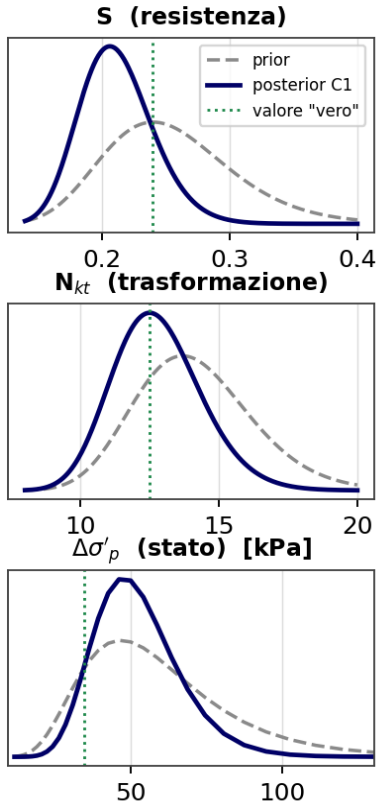
CICLO 1 — in sito

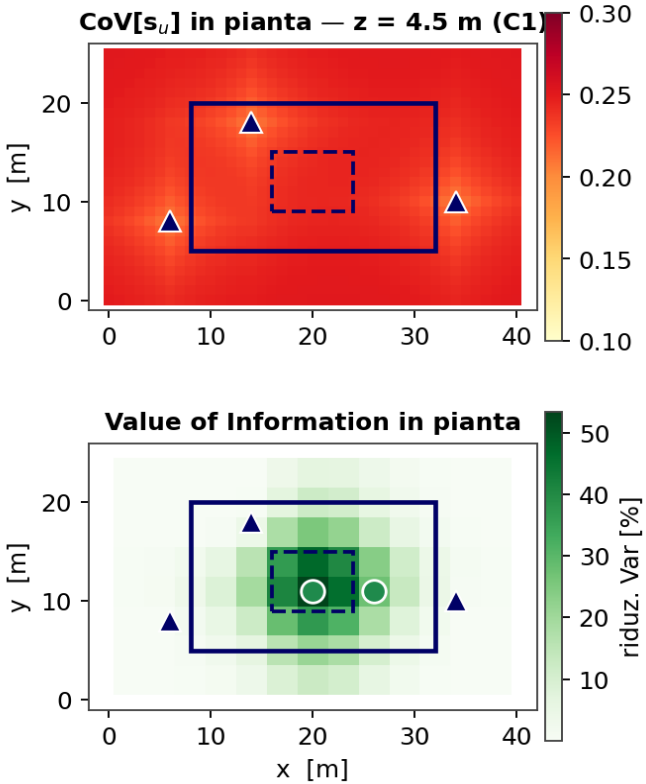
- **3 CPTU continue fino a 14 m, distribuite in pianta** (q_t, f_s, u_2): profilo continuo, alta risoluzione verticale
- **Pianificate su** modello geologico, stati limite e geometria dell'opera, prima di ogni misura

Caso di studio — Ciclo 1: l'aggiornamento

Tre verticali CPT interpretate col modello physics-informed: distribuzioni posteriori di S, N_{kt} e $\Delta\sigma'_p$, banda predittiva sul profilo, mappa del CoV e mappa VoI, entrambe in pianta alla quota critica per l'ULS. La spazializzazione è condotta mediante kriging Bayesiano







SLS prossimo alla soglia di tollerabilità; ULS oltre soglia. La mappa VoI individua in pianta le ubicazioni ottimali delle verticali del ciclo 2.

Caso di studio — Ciclo 2

Perché un secondo ciclo

- Esito del ciclo 1: l'incertezza sul volume mobilitato dall'ULS resta oltre la soglia di tollerabilità, mentre l'SLS è già rientrato — la caratterizzazione non può chiudersi
- La mappa VoI individua le ubicazioni a massima riduzione attesa di varianza per unità di costo, entro la zona critica per l'ULS: l'integrazione della campagna è decisa dai dati, non a priori

Le indagini aggiunte

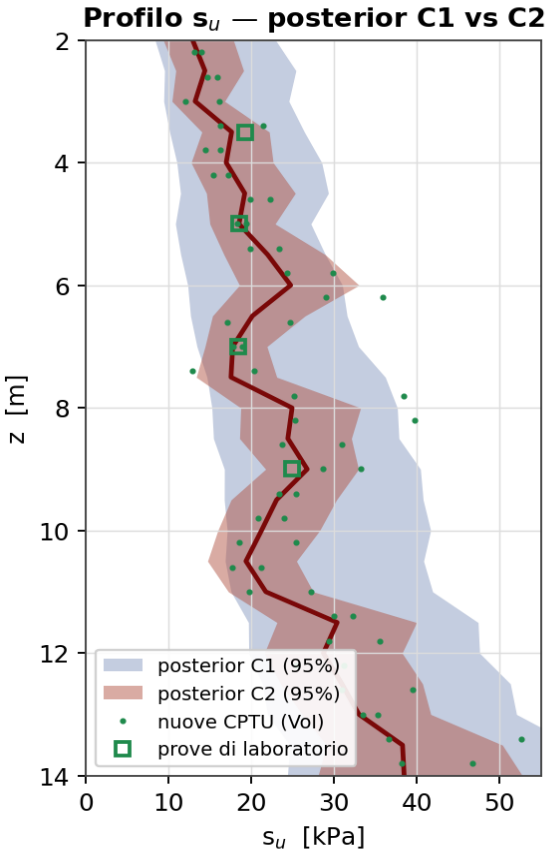
- In sito: 2 CPTU ubicate dalla mappa VoI; 1 sondaggio (S1) con campionamento indisturbato — 4 campioni tra 3.5 e 9 m
- In laboratorio: 4 prove di taglio non drenate (TX–CIU) → s_u; 3 prove edometriche (IL) → σ'_p; classificazione e caratteristiche fisiche (w, LL, IP, γ)
- La posterior del ciclo 1 diventa la prior del ciclo 2: aggiornamento sequenziale, ogni informazione acquisita resta patrimonio del modello

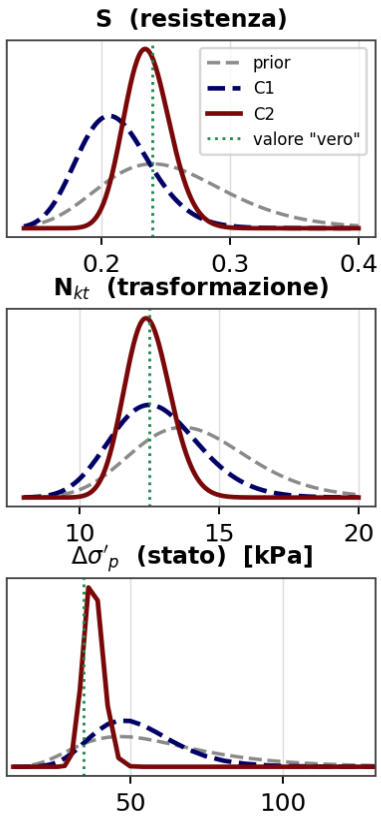
Riduzione della varianza ed esito del ciclo

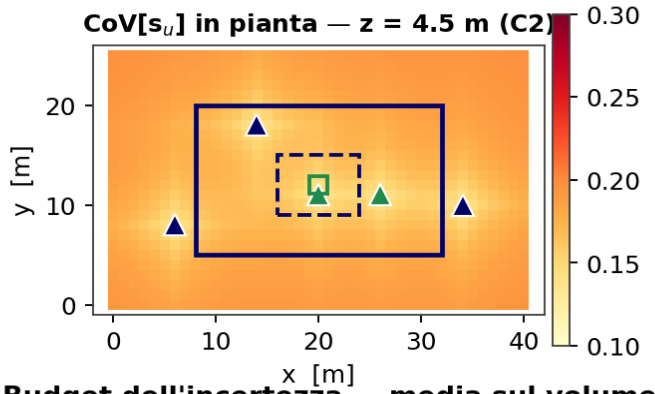
- La stima puntuale sovrastima la dispersione rilevante: ciò che conta è l'incertezza sul volume mobilitato dallo specifico stato limite
- La variabilità intrinseca si riduce per mediazione spaziale; le componenti epistemiche (misura, statistica, trasformazione) si riducono solo migliorando dati e modelli
- Distribuzioni dei parametri chiave strette e ricentrate anche nelle zone critiche per l'ULS; incertezza di trasformazione ridotta dall'aggiornamento sito-specifico dei modelli

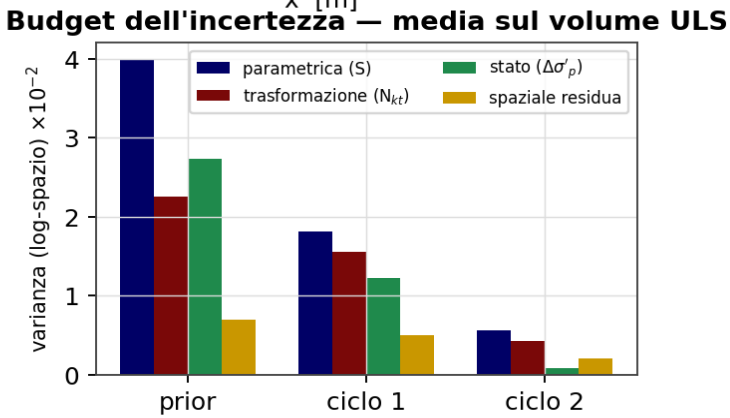
Caso di studio — Ciclo 2: risultati e budget

Due nuove CPT (ubicate dal VoI) e prove di laboratorio: le prove di taglio separano S da N_{kt}, le prove edometriche vincolano $\Delta\sigma'_p$; il budget di varianza quantifica la riduzione delle singole componenti.





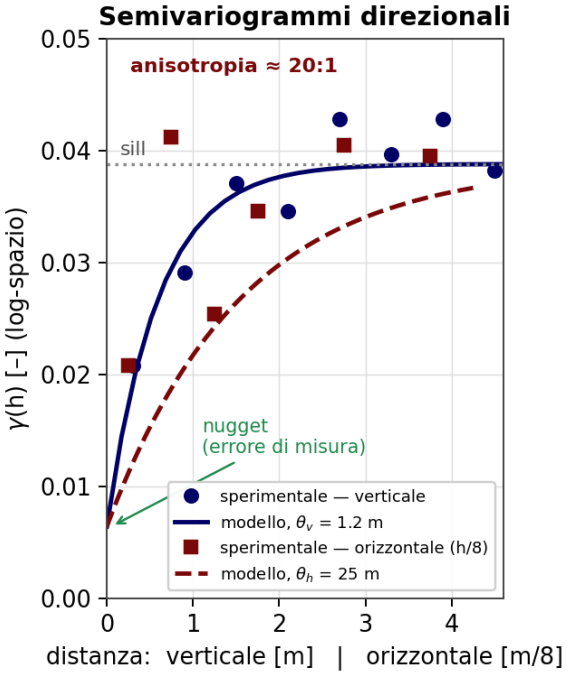


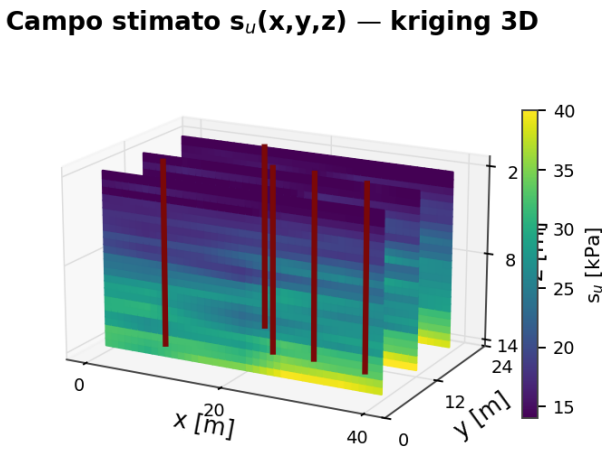


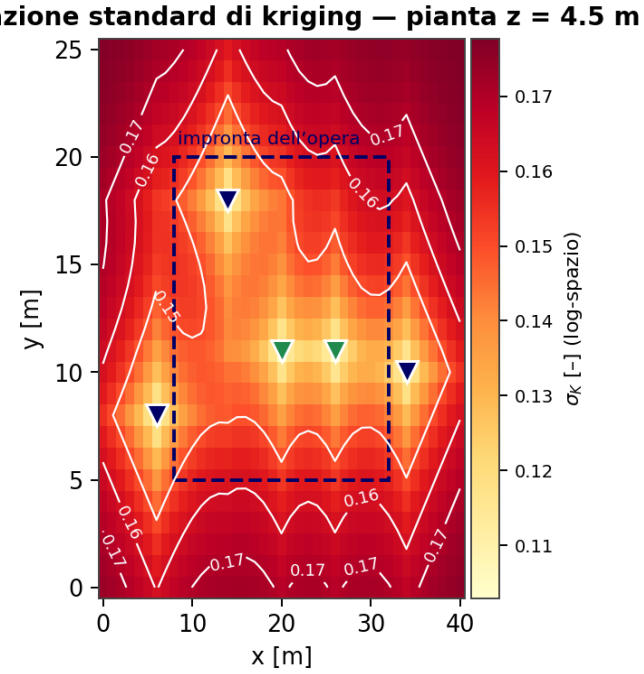
Le componenti epistemiche si riducono con dati e modelli aggiuntivi; il budget di varianza ne quantifica i contributi.

Caso di studio — La spazializzazione: kriging 3D

La spazializzazione del caso di studio, in tre dimensioni: le cinque verticali della campagna completata. La struttura di correlazione è stimata dai residui rispetto al trend fisico: fortemente anisotropa, con continuità orizzontale molto maggiore di quella verticale — conseguenza diretta della genesi deposizionale.



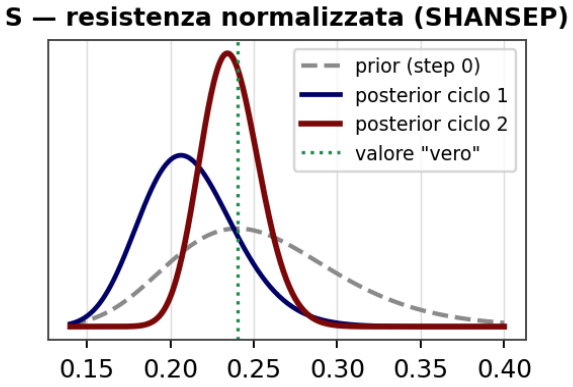


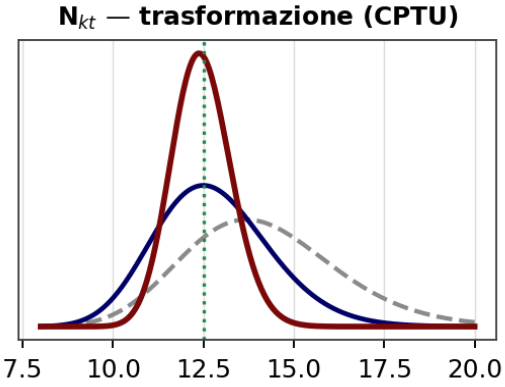


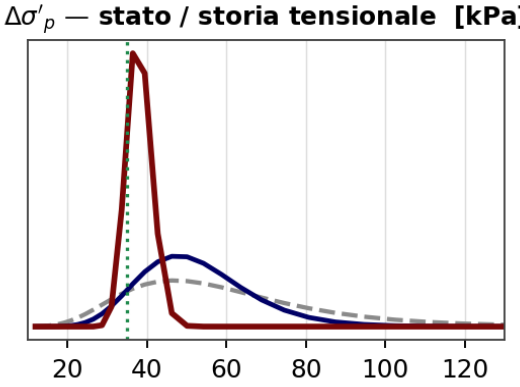
La varianza di kriging è minima in corrispondenza delle verticali e massima negli interspazi: è la mappa quantitativa di ciò che il modello non può stimare a causa dell'ubicazione delle indagini, e costituisce la base dell'analisi VoI per l'ubicazione delle indagini successive.

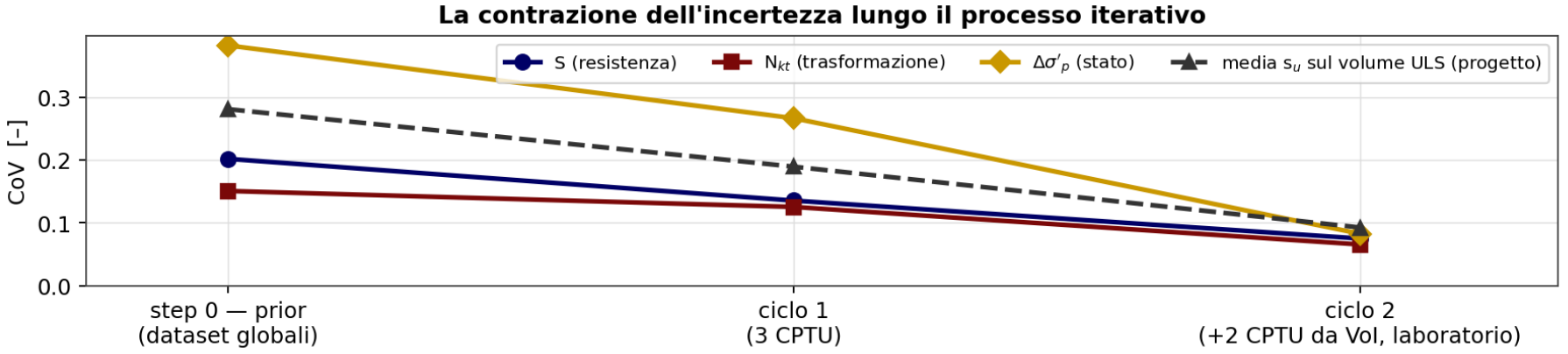
Caso di studio — L'incertezza lungo il processo

Tre parametri di natura diversa — fisico (S), di trasformazione (N_{kt}), di stato tensionale ($\Delta\sigma'_p$, la sovraconsolidazione della crosta) — aggiornati congiuntamente: a ogni ciclo le distribuzioni si contraggono, e con esse l'incertezza sulla grandezza progettuale.









La riduzione iterativa dell'incertezza è misurabile per ciascuno dei parametri multivariati. Il criterio di arresto è dato da soglie di tollerabilità progettuale.

Caso di studio — Dalle posterior ai valori di progetto

Quadro normativo vigente (NTC2018 § 6.2.2; EN 1997-1 § 2.4.5.2)

- Il valore caratteristico è definito come stima ragionata e cautelativa del valore del parametro nello stato limite considerato: la distribuzione a posteriori rende questa definizione quantitativa, riproducibile e tracciabile
- SLS – cedimenti (meccanismo non locale): stima cauta del valor medio sul volume significativo, con riduzione della varianza per mediazione spaziale
- ULS – capacità portante (meccanismo locale): frattile inferiore (5%) letto sulla distribuzione completa — che rivela asimmetrie che l'ipotesi gaussiana nasconderebbe, con esiti non conservativi
- Verifica semiprobabilistica: $X_d = X_k/\gamma_M$ e approccio progettuale invariati; ciò che cambia, e diventa documentabile, è il percorso che conduce a X_k — frattile, supporto spaziale e ipotesi distributiva dichiarati

Evoluzione normativa (EC7-2G)

- La distinzione tra i due meccanismi viene codificata nei valori rappresentativi: valore medio sul volume significativo (Caso A) o frattile 5% per meccanismi locali (Caso B); un quantile intermedio valore nominale (Caso C) resta per scenari intermedi
- Apertura all'analisi pienamente probabilistica (RBD): affidabilità-obiettivo esplicita, con probabilità di superamento dello stato limite calcolata direttamente sulle distribuzioni a posteriori

Criterio di arresto della caratterizzazione

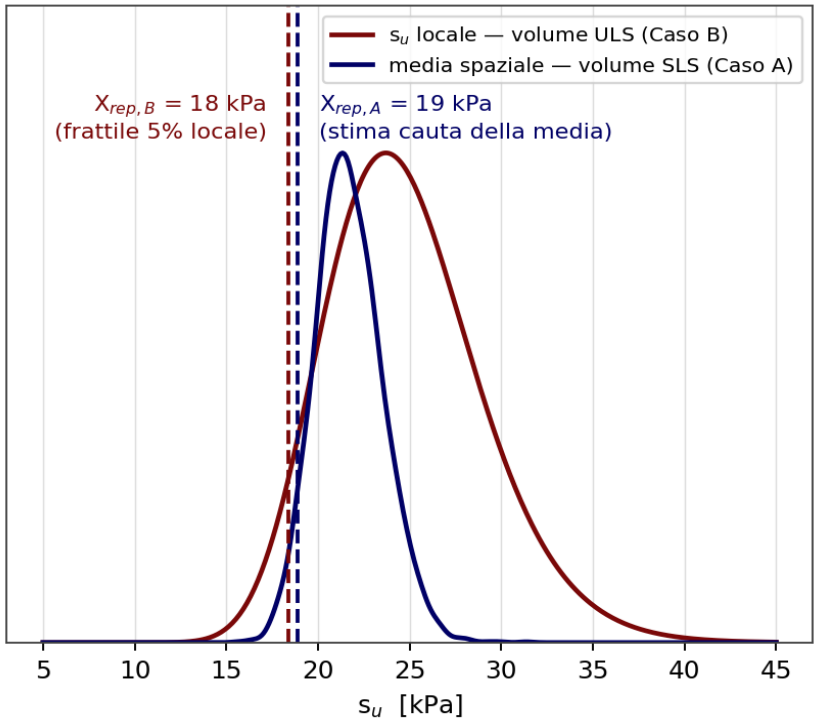
- Incertezza progettuale entro il livello tollerabile per ciascuno stato limite → la caratterizzazione si chiude; altrimenti, nuovo ciclo di indagini

Il ciclo prosegue in esercizio: i dati di monitoraggio sono la likelihood del prossimo aggiornamento (approccio osservazionale Bayesiano).

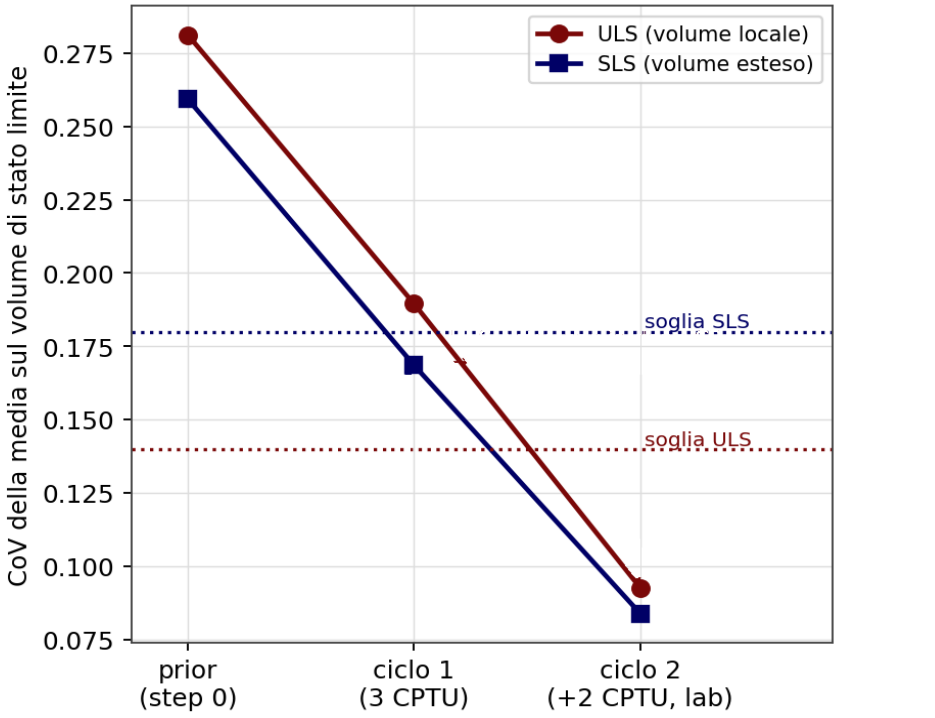
Caso di studio — Risultati finali e arresto

Dalle distribuzioni posteriori ai valori caratteristici per ciascuno stato limite (EC7-2G: Casi A/B) e alla chiusura documentata della caratterizzazione.

Valori rappresentativi per stato limite (EC7-2G)



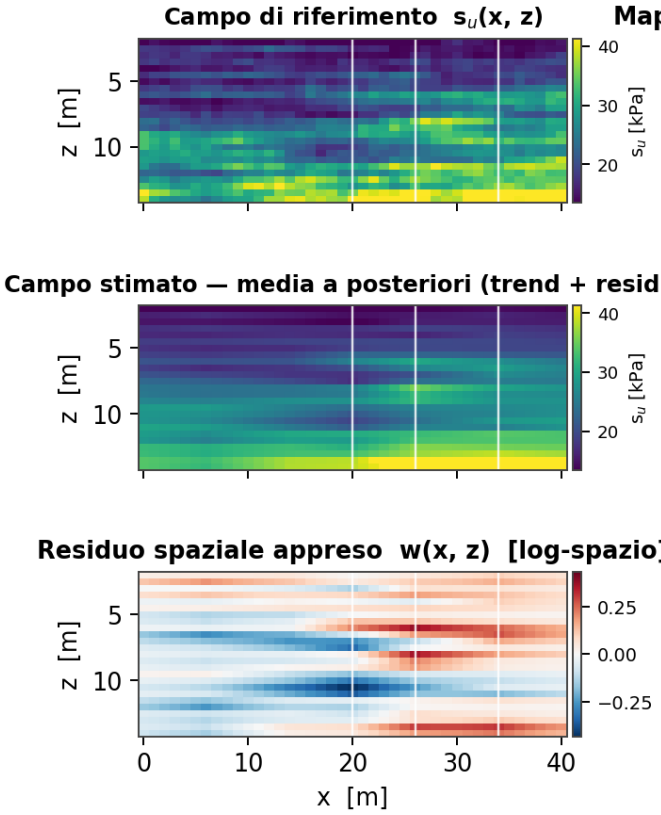
Criterio di arresto — incertezza vs soglia, per stato limite

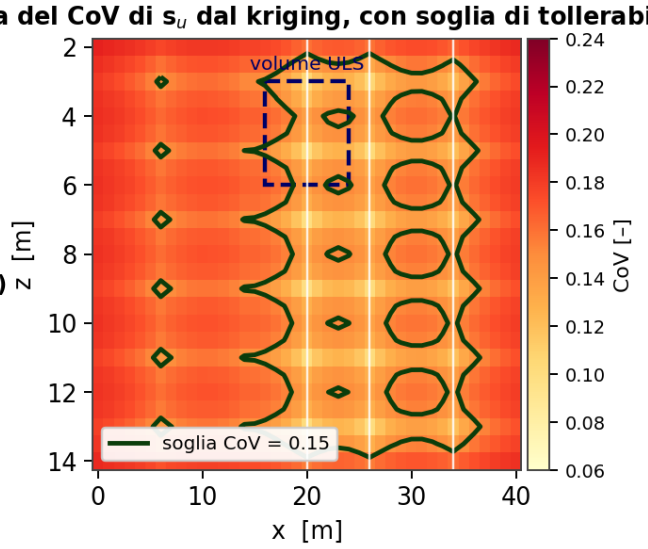


SLS entro soglia dopo il ciclo 1, ULS dopo il ciclo 2: incertezza inferiore alla soglia per entrambi gli stati limite, la caratterizzazione si arresta in modo tracciabile.

Caso di studio — Verifica sul campo di riferimento

Il caso sintetico consente la verifica della calibrazione. La sezione passa per le verticali del ciclo 2; il residuo spaziale appreso evidenzia la struttura ricostruita dal kriging, la mappa del CoV quantifica l'incertezza residua rispetto alla soglia di tollerabilità adottata (CoV = 0.15).

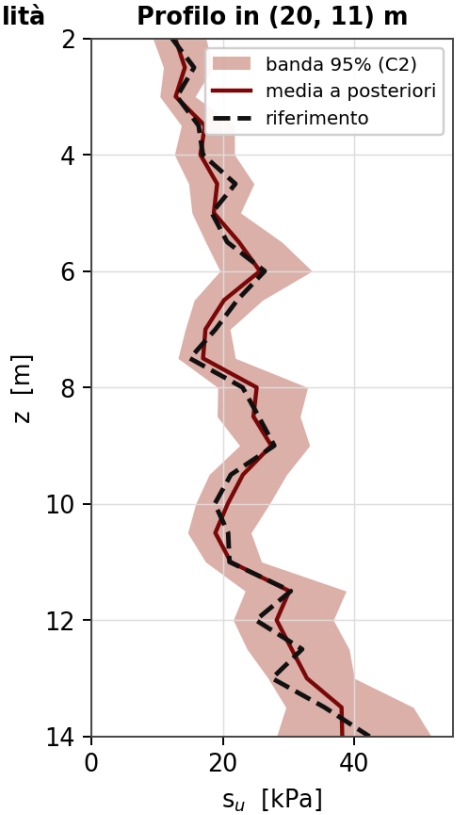




Dal kriging al CoV

$$\ln s_u \sim N(m_K(\mathbf{x}),\ \sigma_K^2(\mathbf{x}))$$

$$\mu = e^{m_K + \sigma_K^2/2} \qquad \mathrm{CoV} = \sqrt{e^{\sigma_K^2} - 1}$$



La banda al 95% copre il campo di riferimento nel 99% dei punti del dominio spaziale. Il criterio di tollerabilità è espresso in termini di CoV, ottenuto da media e varianza di kriging: nel volume ULS il CoV medio è 0.14 e il 68% dei punti rientra nella soglia.

Implementazione: linguaggio, strumenti, requisiti

- **Linguaggio: Python** — open source e gratuito, standard della comunità scientifica (alternative: R, MATLAB, Julia)
- **Librerie di base**: numpy/scipy per algebra e statistica, matplotlib per le figure (tutte risorse gratuite!)
- **Requisiti hardware** — il caso di studio, inferenza e figure comprese, si esegue in meno di 3 s su singolo core; nessun servizio cloud necessario

estratto dal codice del caso di studio (Python + numpy/scipy)

```
import numpy as np;  from scipy.stats import norm, multivariate_normal as mvn

# 1 - dati della campagna e modello fisico (SHANSEP)
P, y, H, R = leggi_campagna(cptu, laboratorio)     # 154 misure
sv0   = 10.0 + 8.2 * P[:, 2]                       # tensione efficace
g_ocr = lambda dsp: m * np.log((sv0 + dsp) / sv0)  # termine di stato

# 2 - struttura di correlazione: kriging bayesiano sul residuo
def acf(A, B):
    dxy = dist_pianta(A, B);  dz = dist_vert(A, B)
    return np.exp(-2 * dxy / theta_h - 2 * dz / theta_v)
C  = sig_w**2 * acf(P, P) + R                      # + errore di misura
Ci = np.linalg.inv(C);  M = H @ Sig0 @ H.T + C

# 3 - posterior di Theta = {S, N_kt, dsigma_p} per marginalizzazione
Sg = np.linalg.inv(iSig0 + H.T @ Ci @ H)
for k, dsp in enumerate(griglia_dsp):
    gk       = g_ocr(dsp)
    logev[k] = mvn.logpdf(y, mean=H @ mu0 + gk, cov=M)
    mus[k]   = Sg @ (iSig0 @ mu0 + H.T @ Ci @ (y - gk))
logev += norm.logpdf(np.log(sigma_p_oed), np.log(griglia_dsp), s_oed)
w   = np.exp(logev + log_prior);  w /= w.sum()     # pesi a posteriori
mu  = (w[:, None] * mus).sum(0)                    # E[ln S], E[ln N_kt]

# 4 - campo condizionato (kriging) e mappa del CoV
K   = sig_w**2 * acf(Q, P)                         # Q: nodi di griglia
mw  = K @ Ci @ (y - (H @ mu + g_ocr(dsp_hat)))     # media di kriging
s2  = sig_w**2 - np.einsum('ij,jk,ik->i', K, Ci, K) + Sig[0,0] + var_t
su  = np.exp(np.log(sv0_Q) + t_bar + mu[0] + mw + s2 / 2)
CoV = np.sqrt(np.exp(s2) - 1)                      # criterio di arresto

# 5 - propagazione ai valori rappresentativi (EC7-2G)
su_vol = monte_carlo(mw, Vw, mu, w, Q_uls, n=3000) # medie sul volume
X_rep  = np.quantile(su_vol, 0.05)                 # Caso A / Caso B
```

Tempi di calcolo misurati

PC portatile, singolo core, numpy/scipy — nessun acceleratore

Inferenza bayesiana — ciclo 1 (90 misure)	**35 ms**
Inferenza bayesiana — ciclo 2 (154 misure)	**90 ms**
Kriging 3D su 26 650 nodi di griglia	**34 ms**
Monte Carlo su volume ULS (3 000 realizzazioni)	**7 ms**
Mappa VoI — 104 ubicazioni candidate	**0.11 s**
Ciclo completo, incluse tutte le figure	**< 3 s**

Il costo computazionale non è il vincolo: l'onere resta nella formulazione del modello e nella validazione dei risultati da parte dell'ingegnere.

Considerazioni riassuntive: algoritmo, dati, ruoli

Ruolo dell'algoritmo composito "ML + geotecnica + Bayes"

- Elaborazione di dati multivariati; spazializzazione; propagazione e aggiornamento delle incertezze; suggerimenti per la pianificazione delle indagini

L'importanza di un'adeguata acquisizione dei dati

- La qualità del risultato finale è determinata almeno tanto dalla campagna di indagini (quantità, qualità, ubicazione delle prove) quanto dagli algoritmi
- L'approccio presentato adotta un metodo "ottimizzante" in relazione ai dati disponibili, ma con dati scarsi o di bassa qualità la distribuzione "posterior" resta molto dispersa
- Quindi, la pianificazione razionale della campagna di acquisizione resta una componente insostituibile della "buona pratica" geotecnica.

L'importanza di un'adeguata interpretazione dei dati

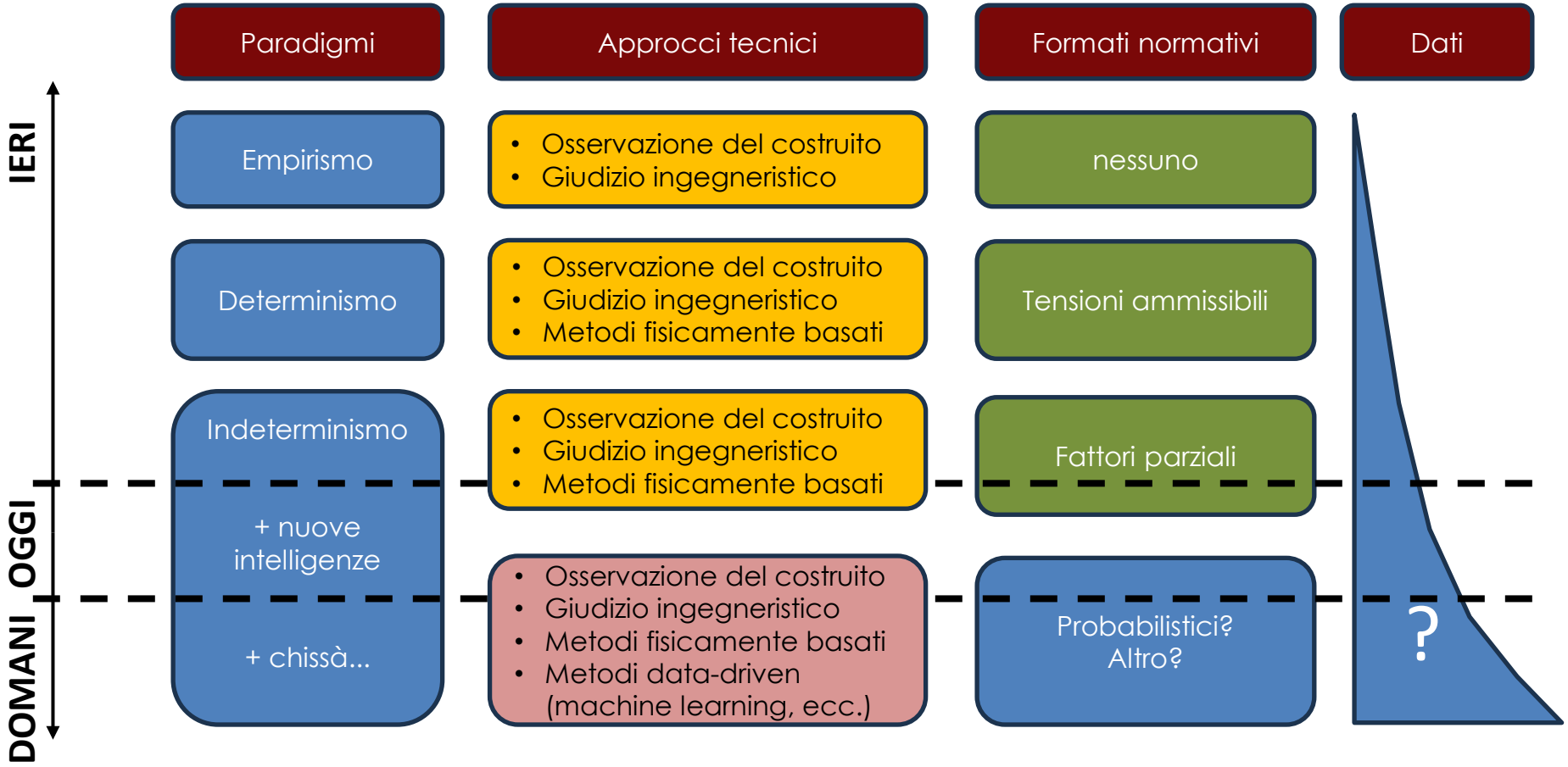
- Trend fisicamente vincolati, incertezze dichiarate e aggiornamento tracciabile rendono l'interpretazione ripetibile e difendibile: la sua qualità non è surrogabile dalla potenza di calcolo

Il ruolo dell'ingegnere geotecnico

- Unità omogenee, stati limite, volumi significativi, soglie di tollerabilità e validazione dei risultati restano decisioni dell'ingegnere: l'algoritmo fornisce stime e suggerimenti, non decisioni

Considerazioni riassuntive: traiettorie future e ruolo dell'ingegnere

Non ci sono discontinuità rispetto alle traiettorie evolutive del passato. La geotecnica, come tutte le discipline ingegneristiche, si è affidata in misura crescente ai dati per affrontare in maniera sempre più esplicita la crescente consapevolezza circa la rilevanza delle incertezze. **Anche il giudizio ingegneristico deve cambiare, in base alle risorse e alle conoscenze disponibili, per evitare "obsolescenza"!**



Funzione del giudizio ingegneristico

Integrare dati, tecnologia e conoscenza geotecnica, senza sostituirsi ad essi.

Fase del processo geotecnico	Buona pratica	Cattiva pratica
Analisi preliminare del sito	Il giudizio costruisce il modello concettuale del sottosuolo dal contesto geologico e morfologico, come ipotesi falsificabile da mettere alla prova con i dati.	Fase saltata o surrogata da analogie acritiche; il modello è dato per scontato, così che i dati non possano né confermarlo né smentirlo.
Definizione degli stati limite rilevanti	Stati limite e situazioni di progetto definiscono grandezze, meccanismi e condizioni di drenaggio: la caratterizzazione è progettata a ritroso da essi.	Si caratterizza "il terreno in sé", a prescindere dai meccanismi; gli stati limite intervengono solo a valle, invertendo la logica del processo.
Definizione dei volumi significativi	La previsione e la verifica iterativa del volume mobilitato dal meccanismo di stato limite definisce la rilevanza della variabilità spaziale e supporta la stima del coefficiente di riduzione della varianza.	Supporto ignorato: valori puntuali o medie su domini arbitrari; la riduzione di varianza per mediazione è trascurata o applicata a volumi non pertinenti.
Pianificazione e svolgimento di campagne di indagine	Indagine disegnata su ipotesi e meccanismi; prove e ubicazioni scelte per stimare media, variabilità e struttura spaziale delle grandezze rilevanti.	Indagine "a ricettario", scollegata da stati limite e volumi; numerosità dettata dal budget; variabilità e scala di correlazione non stimabili.
Definizione di unità geotecniche omogenee	La suddivisione integra evidenza geologica e verifica di stazionarietà: le unità sono ipotesi di omogeneità testabili, non imposte.	Unità tracciate per sola comodità litologica o per clustering automatico; omogeneità postulata e mai verificata, mascherando trend e variabilità interni.
Spazializzazione di dati misurati	Si spazializzano i "valori misurati" (krige-then-transform) con covarianza stimata e dichiarata; si usano stima ed errore, esplicitando l'ipotesi sul campo.	Interpolazione deterministica acritica; incertezza di stima ignorata; σ^2_K letta come distribuzione senza dichiarare l'ipotesi sottostante.
Derivazione di parametri geotecnici	Modelli di trasformazione scelti con giudizio e applicati con la loro incertezza σ^2_T; si propaga l'intera distribuzione, separando incertezza spaziale e di trasformazione.	Correlazioni usate fuori dal dominio di validità come leggi esatte; trasformato il solo valor medio; σ^2_T trascurata; il derivato spacciato per dato.
Assegnazione di valori progettuali	Valore caratteristico come stima cauta coerente con lo stato limite (frattile sul volume pertinente); rispetto della gerarchia misurati→derivati→rappresentativi→progettuali.	Caratteristico scelto a sentimento, scollegato da frattile, supporto e meccanismo; coefficienti parziali applicati meccanicamente; gerarchia dei valori saltata.
Verifica ingegneristica	La verifica è coerente con lo stato limite di partenza. La comunicazione quantitativa dell'incertezza è parte integrante del risultato.	La verifica è deterministica e priva di sensibilità; ipotesi a monte dimenticate; falsa precisione.

Sintesi conclusiva

Rev. Thomas Bayes
(1702-1761)

Grazie

- La traiettoria della disciplina geotecnica verte su paradigmi non deterministici e dato-centrici, "+ nuove intelligenze" — senza discontinuità con il passato. Anche il giudizio ingegneristico deve evolvere, come ha sempre fatto
- AI e ML possono costituire strumenti potenti in un contesto ingegneristico e nel rispetto della Normativa tecnica: se physics-informed e inseriti in un quadro Bayesiano, essi risultano interpretabili, trasferibili con aggiornamento sito-specifico e capaci di quantificare l'incertezza, vista la natura "tipica" dei dati geotecnici.
- La qualità e la consistenza della campagna di indagini restano imprescindibili: pianificata su opera, stati limite e modello geologico, e integrata in modo mirato sulla base del valore atteso dell'informazione. Le incertezze si accumulano lungo il processo e quelle introdotte in acquisizione non sono eliminabili a valle: la fase di acquisizione condiziona irreversibilmente interpretazione e progettazione
- Il fine del processo non deterministico è ridurre l'incertezza progettuale entro livelli tollerabili per ciascuno stato limite, e mostrare se necessario il processo iterativo
- L'ingegnere geotecnico resta al centro: l'AI ne rafforza il giudizio, non lo sostituisce!