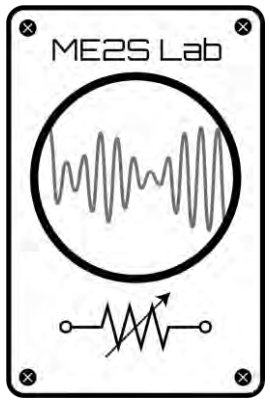


AI e Misure:
il ruolo dell'incertezza

Gianluca Ciattaglia, Susanna Spinsante
Laboratorio di Misure Elettriche, Elettroniche e Sensori (ME2S Lab)
Dipartimento di Ingegneria dell'Informazione
29.10.2026

Sommario

- Contesto
- La misura e l’incertezza nella metrologia
- L’AI nelle misure
- L’incertezza dell’AI
- Quantificare l’incertezza
  - Nella metrologia
  - Nell’AI
- Una nuova tassonomia
- Casi applicativi
- Conclusioni

Il contesto

- L'intelligenza artificiale sta trasformando profondamente quasi tutti i domini della conoscenza, dalla medicina all'ingegneria, dalla finanza alle scienze sociali, offrendo nuove capacità di analisi, previsione e supporto decisionale.
- La sua crescente integrazione nei processi di ricerca, sviluppo e misura sta ridefinendo metodi, strumenti e paradigmi consolidati.
- Ma ci chiediamo:
  - L'intelligenza artificiale sta rivoluzionando la metrologia, ma ogni misura ha davvero bisogno di un modello AI per essere migliore?
  - Tra sensori, modelli fisici e algoritmi di AI, quale componente aggiunge realmente informazione e quale soltanto complessità?
  - **Si può trascurare il contributo di incertezza portato dall'AI applicata alle misure?**

La misura e l'incertezza

- Incertezza: il termine indica «il dubbio» circa il risultato di misura di una grandezza fisica

Measurement *is one of the most common processes in human activity.* ***Measured values of quantities*** *are required for a diverse range of applications and, for each of these values, a statement is needed about its* ***credibility****. Such a statement is usually expressed in terms of* ***measurement uncertainty****.*

The ***two*** *components, the measured* ***value*** *and the associated* ***uncertainty****,* ***together*** *constitute the most common way of reporting a* ***measurement result.***

Joint Committee for Guides in Metrology (JCGM), Guide to the expression of uncertainty in measurement — Part 1: Introduction, JCGM GUM-1:2023

L’incertezza secondo GUM e VIM

*When reporting the result of a **measurement of a physical quantity**, it is **obligatory** that some **quantitative indication** of the quality of the result be given so that those who use it can assess its **reliability**. Without such an indication, measurement results cannot be compared, either among themselves or with reference values given in a specification or standard.*

UNCERTAINTY (of measurement)

“parameter, associated with the result of a measurement, that characterizes the **dispersion** of the values that could reasonably be attributed to the measurand”

GUM fornisce una metodologia chiara:

- si ripete la misura molte volte
- si ottengono risultati leggermente diversi
- la dispersione di tali risultati rappresenta l'incertezza
- più grande è la dispersione, maggiore è l'incertezza

JCGM 100:2008 **GUM** 1995 with minor corrections Evaluation of measurement data — **Guide to the expression of Uncertainty in Measurement**
JCGM-WG2-CD-01 International Vocabulary of Metrology (**VIM**) Fourth edition, 11 January 2021

L’AI nelle misure - I

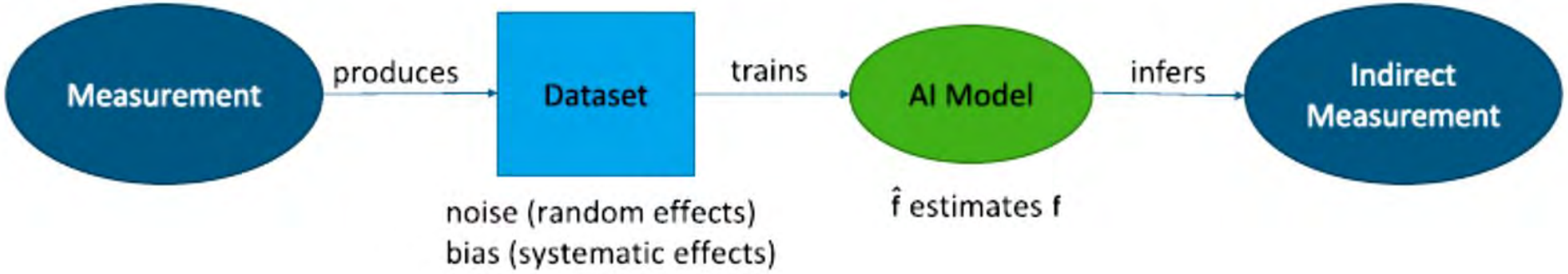
- Misure indirette
- AI come strumento di elaborazione del dato misurato
  - Dalle acquisizioni dei sensori, l'AI **estrae informazioni**
- AI come modello di misura (soft sensor o virtual sensor)
  - L'AI **sostituisce**, in tutto o in parte, il **modello fisico** che lega osservabili a grandezza di interesse
- AI come componente di un “Synthetic Instrument”
  - Strumento in cui il risultato finale è ottenuto combinando sensori fisici, acquisizione dati, modelli software, algoritmi AI
  - AI è **componente costitutiva dello strumento**

L’AI nelle misure - II

- AI come strumento di elaborazione del dato misurato
  - Esempi: classificazione di immagini provenienti da sistemi di visione, analisi di segnali biomedicali (EEG, ECG, EMG), *riconoscimento di difetti in controlli non distruttivi*, identificazione di eventi o anomalie in dati provenienti da sensori.
  - L'output dell'AI non è necessariamente una misura numerica, ma può essere una classe, una diagnosi o una decisione.
- AI come modello di misura (soft sensor o virtual sensor)
  - $x_1, x_2, \dots, x_n \rightarrow AI \rightarrow y$, dove y è il misurando
  - Esempi: stima di temperatura, pressione o concentrazione chimica da misure indirette più semplici da fare, monitoraggio di processi industriali, sistemi di misura virtuali, predictive sensing.
  - AI è parte integrante della catena di misura.
- AI come componente di un “Synthetic Instrument”
  - La misura finale dipende contemporaneamente da incertezza dei sensori, qualità del dataset, capacità di generalizzazione del modello, condizioni operative

L’importanza dell’incertezza dell’AI

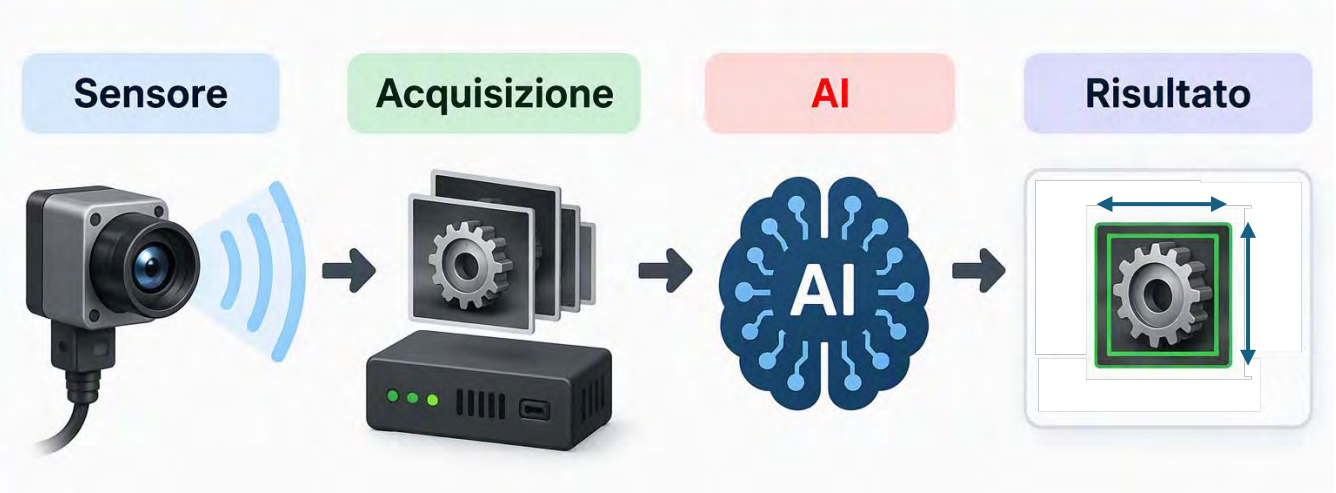
- Anche una misura assistita da AI è soggetta a incertezza
- **Deve essere quantificata** per:
  - capire quando un risultato è affidabile
  - supportare decisioni automatiche o umane
  - aumentare la fiducia nel sistema
  - gestire il rischio derivante da decisioni errate



- Esempio: guida autonoma

AI e misure… e l’incertezza?

Indipendentemente dall'applicazione specifica, l'AI deve essere vista come un **nuovo blocco della catena di misura**:



quando un sistema di misura utilizza l'AI per stimare una grandezza fisica o classificare un fenomeno, **come si quantifica correttamente l'incertezza del risultato?**

AI, misure e incertezza

- La comunità della metrologia utilizza da decenni il framework **GUM** (Guide to the Expression of Uncertainty in Measurement), **invece** la comunità AI ha sviluppato **proprie definizioni e metodi** di uncertainty quantification.
- Spesso le terminologie sono diverse, talvolta incompatibili.
- L'incertezza introdotta dal modello AI deve essere **integrata nel bilancio di incertezza (metrologica)** complessivo dello strumento
- Vengono proposti **due contributi principali** di incertezza:
  - **Aleatoric uncertainty** ← dovuta ai dati e alle misure
  - **Epistemic uncertainty** ← dovuta al modello AI

che insieme concorrono all'incertezza complessiva della misura.

S. Shirmohammadi, F. Wang and C. -H. Hsu, "Review and Performance Evaluation of Uncertainty Quantification in Data-Driven AI-Assisted Measurements," in *IEEE Open Journal of Instrumentation and Measurement*, vol. 4, pp. 1-22, 2025, Art no. 2500722, doi: 10.1109/OJIM.2025.3621742.

L’incertezza secondo la metrologia (GUM, VIM)

Type A: osservazioni **ripetute** (lo stesso misurando viene misurato più volte)

Si calcolano:

- media campionaria;
- varianza sperimentale;
- deviazione standard **della media** σ → **Type A standard uncertainty ($u_A(x)$)**

Assumendo una distribuzione normale:

±1σ → 68.27%
±2σ → 95.45%
±3σ → 99.73%
} confidence level

Type B: osservazioni ripetute non sono disponibili oppure non sono praticabili

Le informazioni possono provenire da:

- dati precedenti, esperienza;
- specifiche del costruttore, certificati di calibrazione;
- letteratura tecnica;
- handbook e reference data

In questo caso l'incertezza **($u_B(x)$)** deriva da una valutazione scientifica informata e non da una statistica sperimentale, ma la si riconduce ad una PDF con propria media e deviazione standard

Incertezza combinata

Quando più sorgenti contribuiscono all'incertezza (ingressi multipli, sensori multipli, errori differenti) si applica la legge di propagazione dell'incertezza (GUM)

Caso sorgenti indipendenti: $u_c = \sqrt{u_1^2 + u_2^2 + \cdots}$ e, in particolare: $u_c = \sqrt{u_A^2 + u_B^2}$

Caso sorgenti non indipendenti (modello **lineare**): $u_C^2 = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} u(x_i, x_j)$

Caso sorgenti non indipendenti (modello **NON lineare**): $u_C^2 = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i) + \sum_{i=1}^{N}\sum_{j=1}^{N} \left[\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)^2 + \frac{\partial f}{\partial x_i}\frac{\partial^3 f}{\partial x_i\, \partial x_j^2}\right] u^2(x_i)u^2(x_j)$

Incertezza estesa: $U = ku_C$ quando occorre fornire un livello di fiducia superiore al 68%

L’incertezza secondo la letteratura AI

L'AI utilizza *Accuracy*, *Precision*, *Recall*, *F1-score* per valutare **i modelli**

- Metriche che descrivono il **comportamento medio del modello** su un **dataset**.
- **Non** ci dicono quanto sia affidabile una **singola predizione**.

Nell’AI, l’incertezza può originare da:

- I **dati**: derivano da misure reali, contengono rumore, errori casuali, errori sistematici residui. Essi vengono trasferiti al dataset.
  - **Aleatoric uncertainty:** omoschedastica (tutti i campioni hanno la stessa incertezza) o eteroschedastica (ogni campione possiede un'incertezza diversa) – caso tipico in applicazioni reali.
- Il **modello**: apprende una funzione approssimata della realtà, tipicamente $\hat{f} \neq f$. Questa approssimazione introduce ulteriore incertezza.
  - **Epistemic uncertainty:** dovuta al modello, e origina da architettura non perfetta, parametri appresi subottimali, dati non rappresentativi, campioni Out-Of-Distribution (OOD).
- Poiché tipicamente coesistono, si deve valutare la **total uncertainty**: $U_{tot} = U_{aleatoric} + U_{epistemic}$ con le *opportune formulazioni statistiche*.

Osservazioni ripetute e incertezza aleatorica

- Caso aleatorico: il modello resta fisso, si varia l'ingresso.
- I modelli AI sono normalmente deterministici, quindi allo stesso input corrisponde lo stesso output → occorre generare artificialmente *osservazioni ripetute*
  - **Monte Carlo Simulation (MCS)**
    - Si modellano gli ingressi tramite distribuzioni statistiche: rumore, tolleranze, variabilità ambientale.
    - Si campionano molte realizzazioni dell'ingresso e si osserva come varia l'uscita.
  - **Test-Time Augmentation (TTA)**
    - Molto usata per immagini.
    - Durante l'inferenza si generano versioni modificate dell'input: rotazioni, flip, crop, variazioni di luminosità.
    - Le predizioni risultanti costituiscono osservazioni ripetute.

Osservazioni ripetute e incertezza epistemica

- Caso epistemico: si modifica il modello lasciando fisso l'input.
  - **Tecniche Ensemble**: si utilizzano più modelli, e ogni modello produce una propria predizione. L'insieme delle predizioni costituisce il campione statistico per stimare l'incertezza.
  - **Bayesian Neural Networks (BNN)**: i pesi vengono rappresentati come distribuzioni di probabilità, pertanto pur in presenza dello stesso input, con pesi campionati diversamente si ottengono output differenti
  - **Monte Carlo Dropout (MCD)**: durantel'inferenza alcuni neuroni vengono disattivati casualmente, si eseguono molte forward pass e si ottiene una distribuzione di output
  - **Variational Inference**: approssima il risultato bayesiano tramite distribuzioni più semplici (tipicamente gaussiane), permette di costruire la distribuzione dell'output.
  - **Bayesian Decision Trees**: versione bayesiana degli alberi decisionali, in cui struttura dell'albero, parametri e split diventano variabili casuali.

Metodi di quantificazione dell'incertezza

- Si considerano regressione e classificazione
- **Regressione**: il modello predice un valore numerico continuo
  - Secondo VIM è una misura indiretta
- Valutazione dell'incertezza:

EPISTEMICA	**Type A:**	uno dei metodi visti per generare osservazioni ripetute, poi si prende la deviazione standard sperimentale della media delle predizioni
	Type B:	osservazioni ripetute non sono possibili: si valuta l'incertezza del modello mediante metodi basati sul gradiente che indica quanto l'output predetto varia in risposta a piccole perturbazioni dell'input o delle sue caratteristiche
ALEATORICA	**Type A:**	uno dei metodi visti per generare osservazioni ripetute e poi si prende la deviazione standard sperimentale della media delle predizioni
	Type B:	stimare l'incertezza dell'output dovuta all'incertezza dell'input con propagazione analitica dell'incertezza (a seconda che $\hat{f}$ sia lineare o meno): *Gaussian Likelihood-Based Loss, Quantile Regression, Mixture Density Networks*

Metodi di quantificazione dell'incertezza

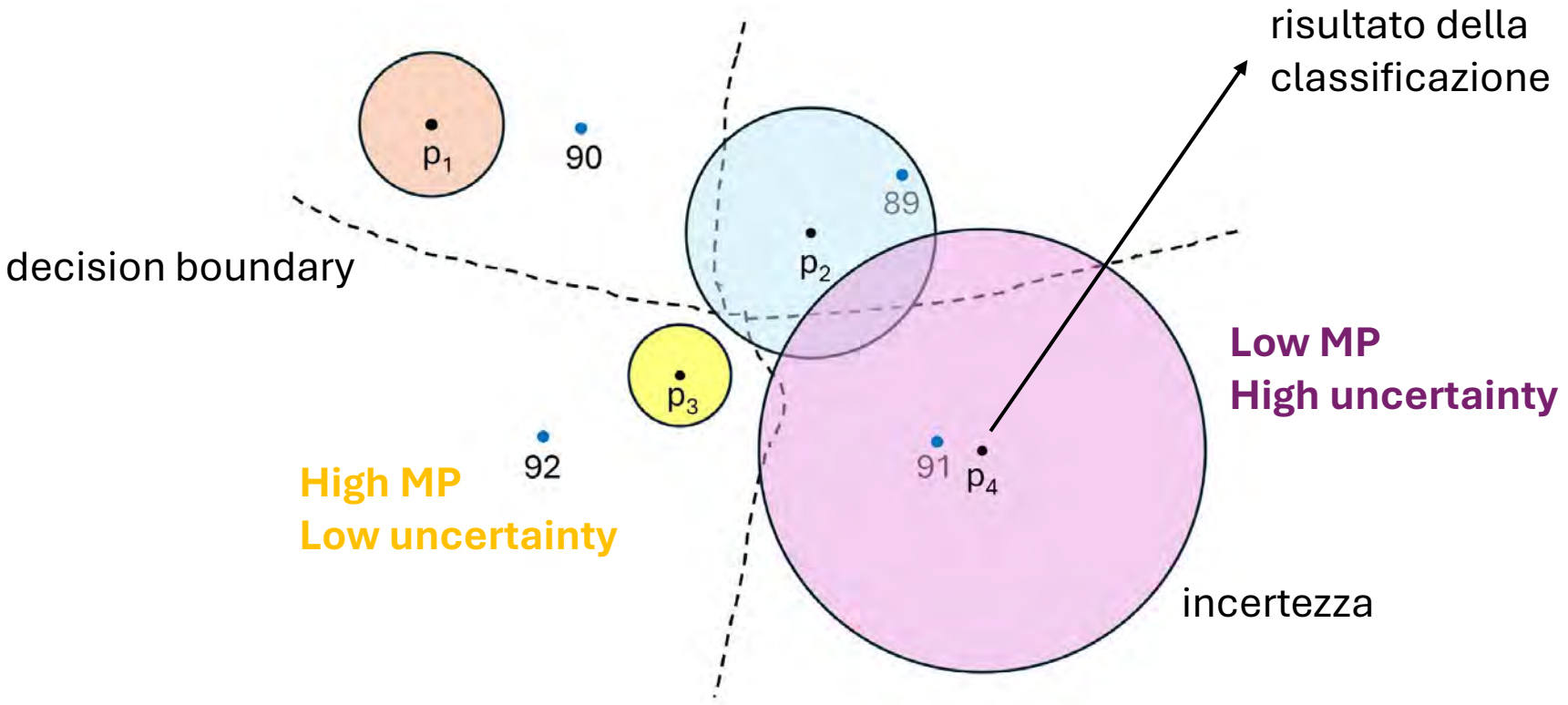
- Classificazione: il modello classifica un dato input in una **classe** predefinita in funzione delle sue caratteristiche descrittive (feature).
- E' una misura?
  - La classe è una quantità ordinale: secondo VIM è una misura indiretta, manca una procedura GUM per la stima di incertezza
  - La classe è una proprietà nominale: secondo VIM non è una misura ma secondo altri standard (vocabulary of the International Federation of Clinical Chemistry and Laboratory Medicine) è definita *examination* quindi richiede **incertezza** (**dubbio** circa la classificazione) → da **non confondere** con la **Misclassification Probability** (MP)
- Valutazione dell'incertezza:

EPISTEMICA	**Type A:**	uno dei metodi visti per generare le osservazioni ripetute (serie di probabilità di previsione) e poi si applica uno tra i metodi Deviation of Predicted Probabilities, Entropy of the Mean, Mean of the Entropy, Maximum Entropy alle probabilità predittive
	Type B:	come per la regressione
ALEATORICA	**Type A:**	uno dei metodi visti per generare le osservazioni ripetute, poi stessi metodi della epistemica Type A applicati alle probabilità predittive
	Type B:	propagazione analitica dell'incertezza (a seconda che $\hat{f}$ sia lineare o meno), entropia delle probabilità predittive risultanti: $u = H(R) = -\sum_{i=1}^{R} p_i \log_b(p_i)$ dove H è la funzione di entropia, R è il numero delle classi, p_i la probabilità predittiva della classe i-esima, e si prende $b = R$ per avere un risultato compreso tra 0 e 1

G. Nordin et al., "Vocabulary On nominal property, examination, and related concepts for clinical laboratory sciences (IFCCIUPAC recommendations 2017)," Pure Appl. Chem., vol. 90, no. 5, pp. 913–935, 2018

Incertezza e Misclassification Probability (MP)

Non sono la stessa cosa!



F. Wang, S. Shirmohammadi and C. -H. Hsu, "On the False Notion of Misclassification Probability as Uncertainty in Machine Learning-Assisted Measurements," *2025 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Chemnitz, Germany, 2025, pp. 1-5, doi: 10.1109/I2MTC62753.2025.11078994.

S. Shirmohammadi, F. Wang and C. -H. Hsu, "Review and Performance Evaluation of Uncertainty Quantification in Data-Driven AI-Assisted Measurements," in *IEEE Open Journal of Instrumentation and Measurement*, vol. 4, pp. 1-22, 2025, Art no. 2500722, doi: 10.1109/OJIM.2025.3621742.

Una nuova tassonomia

AI Uncertainty		Measurement Uncertainty		
Inference model	Type of uncertainty	Type A evaluation		Type B evaluation
		1- generate multiple observations	2- quantify the uncertainty	
Regression	Epistemic	Ensembles; Bayesian methods such as Bayesian decision trees (for tree-based models) or BNN approximations (for ANN models) such as Monte Carlo dropout or variational inference.	Quantify the uncertainty as the experimental standard deviation of the observations' mean.	Gradient-based methods; pure BNN (currently impractical).
	Aleatoric	Monte Carlo simulation; test-time augmentation.		Analytical propagation of uncertainty; Gaussian likelihood-based loss; Quantile regression; Mixture Density Networks.
Classification	Epistemic	Use the same epistemic Type A methods for regression to generate multiple predictive probabilities for each class.	Quantify the uncertainty by applying DPP, entropy_mean, mean entropy, or max entropy to the predictive probabilities.	Use the same epistemic Type B methods for regression.
	Aleatoric	Use the same aleatoric Type A methods for regression to generate multiple predictive probabilities for each class.		Analytical propagation of uncertainty; entropy of the predicted probabilities.
Large Language Model (LLM)	Epistemic	Prompt ensembles; sampling with different seeds or temperatures; dropout at inference; model ablations.	Quantify the uncertainty by applying entropy at the token level or sequence level to the observations.	Weight-space analysis.
	Aleatoric	Token sampling variability.		Temperature scaling of logits.

LLM → work in progress

S. Shirmohammadi, F. Wang and C. -H. Hsu, "Review and Performance Evaluation of Uncertainty Quantification in Data-Driven AI-Assisted Measurements," in *IEEE Open Journal of Instrumentation and Measurement*, vol. 4, pp. 1-22, 2025, Art no. 2500722, doi: 10.1109/OJIM.2025.3621742.

Collaudo in linea di ventole industriali con radar FMCW a 77 GHz

- **Contesto**: ogni ventola prodotta (EMC FIME) va accettata o scartata senza rallentare la linea;
- **Limiti attuali**: accelerometro sul DUT: montaggio manuale. Sul banco: accoppiamento indiretto, attenuazione. LDV: costoso, ingombrante;
- **Perché il radar:** economico, robusto a polvere e rumore acustico, compatto;
- **Grandezze di collaudo:** rpm (efficienza, portata) e displacement (sbilanciamento, disallineamento), secondo ISO 14695:2003.

Perché l'AI non migliora l'accuratezza

Misura rpm: L'errore è quantizzazione della griglia DFT (Type B, $\Delta f/\sqrt{12}$), non si corregge con un regressore.
Displacement: Il radar misura la media nel lobo e il limite è la definizione del misurando e l'SNR, non il modello.
Modello di misura noto: uno appreso aggiunge incertezza epistemica senza ridurre i contributi dominanti.

Casi di Inquinanti inorganici in acqua: Raman per l'analisi in situ

- **Contesto applicativo**: Nitrati, solfati e nitriti da agricoltura e industria: serve analisi rapida in situ, senza un esperto.
- **Perché il Raman:** Non distruttivo, nessuna preparazione del campione. La posizione del picco identifica l'analita, l'intensità segue la concentrazione.
- **Limiti**: Bassa sensibilità, baseline da fluorescenza, rumore shot, shift del picco per interazioni ioniche in miscela.

Dove l'AI migliora lo strumento

Correzione sistematica: rimuove fluorescenza e deriva della linea di base senza l'intervento di un esperto.

Ripetibilità: estrae l'ampiezza del picco da un modello gaussiano, che resta stabile anche quando il picco si sposta.

Da qualitativa a quantitativa: converte lo spettro in un valore di concentrazione con una taratura lineare, semplice e interpretabile.

Generalizzazione: il sistema è tarato su soluzioni con un solo analita e quantifica anche le miscele.

Principali conclusioni

- L’introduzione dell’AI in sistemi o metodi di misura va attentamente valutata, in relazione al problema che essa può aiutare a risolvere
- L'incertezza nei sistemi di misura assistiti da AI deve essere trattata integrando i concetti della metrologia (GUM) e quelli dell’AI
- La probabilità di misclassificazione non rappresenta una misura di incertezza, ma una grandezza complementare
- L'uso di osservazioni ripetute migliora sensibilmente la quantificazione dell'incertezza, soprattutto nei problemi di classificazione multiclasse, ma non esiste una soluzione universalmente ottimale