

Efficienze POLA-R (NN)

F. Noferini

La stima delle efficienze POLA-R

Dati due SiPM (1,2) su una piastra con:

- ϵ_1 = efficienza assoluta (non condizionata) del SiPM 1
- ϵ_2 = efficienza assoluta (non condizionata) del SiPM 2
- $\epsilon_{1|2}$ = efficienza condizionata di 1 **dato che** 2 ha visto il segnale (cioè probabilità che 1 scatti, sapendo che 2 è scattato)
- $\epsilon_{2|1}$ = efficienza condizionata di 2 **dato che** 1 ha visto il segnale
- ϵ_{AND} = efficienza di coincidenza (AND), ovvero probabilità che **entrambi** scattino insieme.

$$\epsilon_{1|2} = \frac{P(1 \cap 2)}{P(2)} = \frac{\epsilon_{AND}}{\epsilon_2}$$

Quello che
misuriamo

$$\epsilon_{2|1} = \frac{P(1 \cap 2)}{P(1)} = \frac{\epsilon_{AND}}{\epsilon_1}$$

$$\epsilon_{Maj} = \epsilon_{AND}^{12} \epsilon_{AND}^{34} \left[\frac{1}{\epsilon_{1|2}} + \frac{1}{\epsilon_{2|1}} + \frac{1}{\epsilon_{3|4}} + \frac{1}{\epsilon_{4|3}} - 3 \right]$$

Quello che vogliamo
accoppiando due piastre

Modello semplificato (attuale)

Nell'ipotesi di SiPM indipendenti

- $\epsilon_1 = \epsilon_{1|2}$
- $\epsilon_2 = \epsilon_{2|1}$
- $\epsilon_3 = \epsilon_{3|4}$
- $\epsilon_4 = \epsilon_{4|3}$
- $\epsilon_{AND}^{12} = \epsilon_{1|2} \cdot \epsilon_{2|1}$
- $\epsilon_{AND}^{34} = \epsilon_{3|4} \cdot \epsilon_{4|3}$



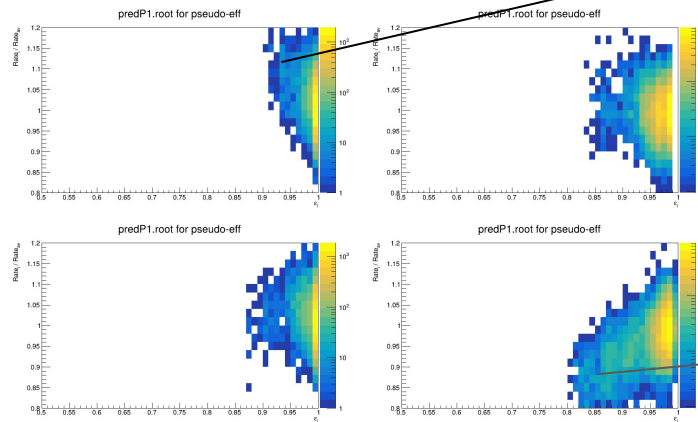
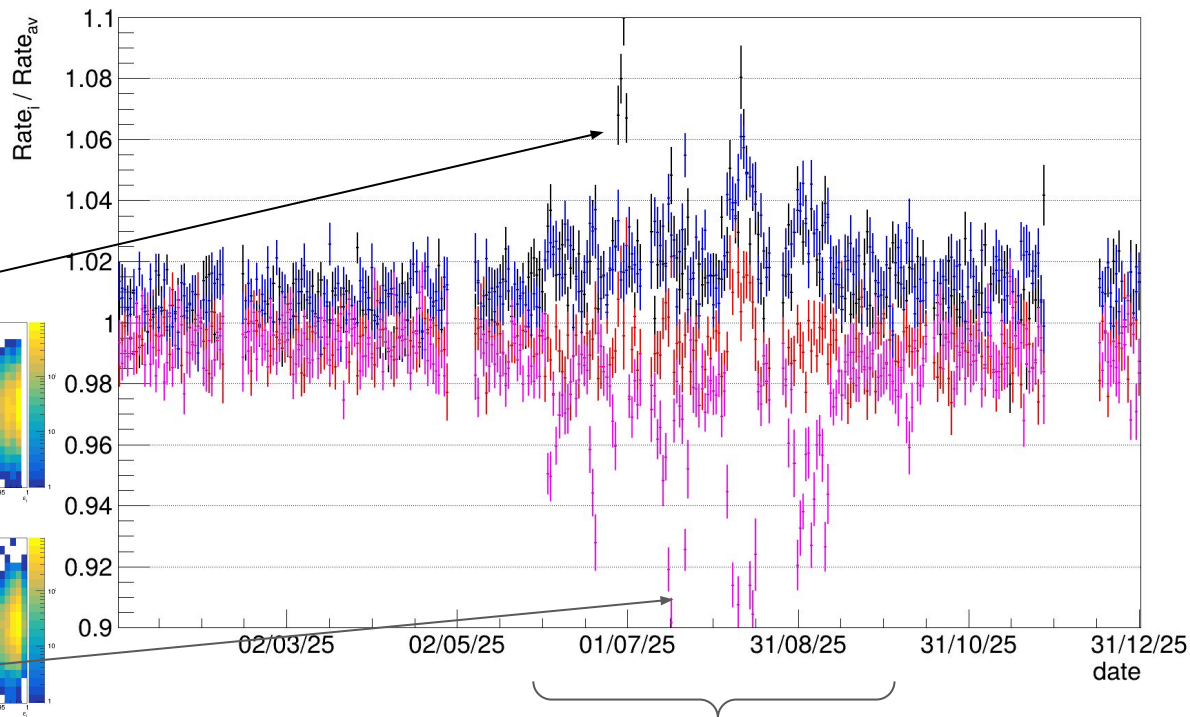
$$\epsilon_{Maj} = \epsilon_1 \epsilon_2 \epsilon_3 \epsilon_4 \left(\frac{1}{\epsilon_1} + \frac{1}{\epsilon_2} + \frac{1}{\epsilon_3} + \frac{1}{\epsilon_4} - 3 \right)$$

E' semplice da calcolare
(pseudo-efficienze)

Qual è l'effetto sistematico introdotto da questa approssimazione? Non facile da stimare
Si può rilasciare questa condizione?

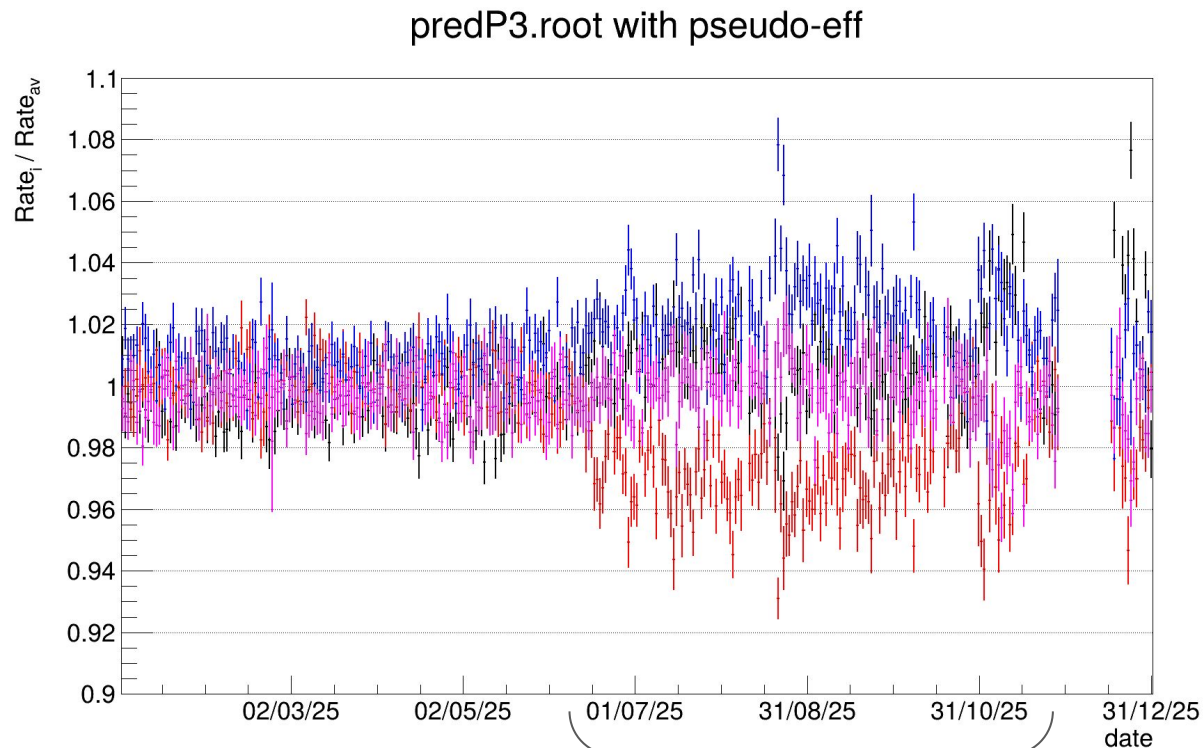
Grafici di controllo POLA-01 (2025)

predP1.root with pseudo-eff



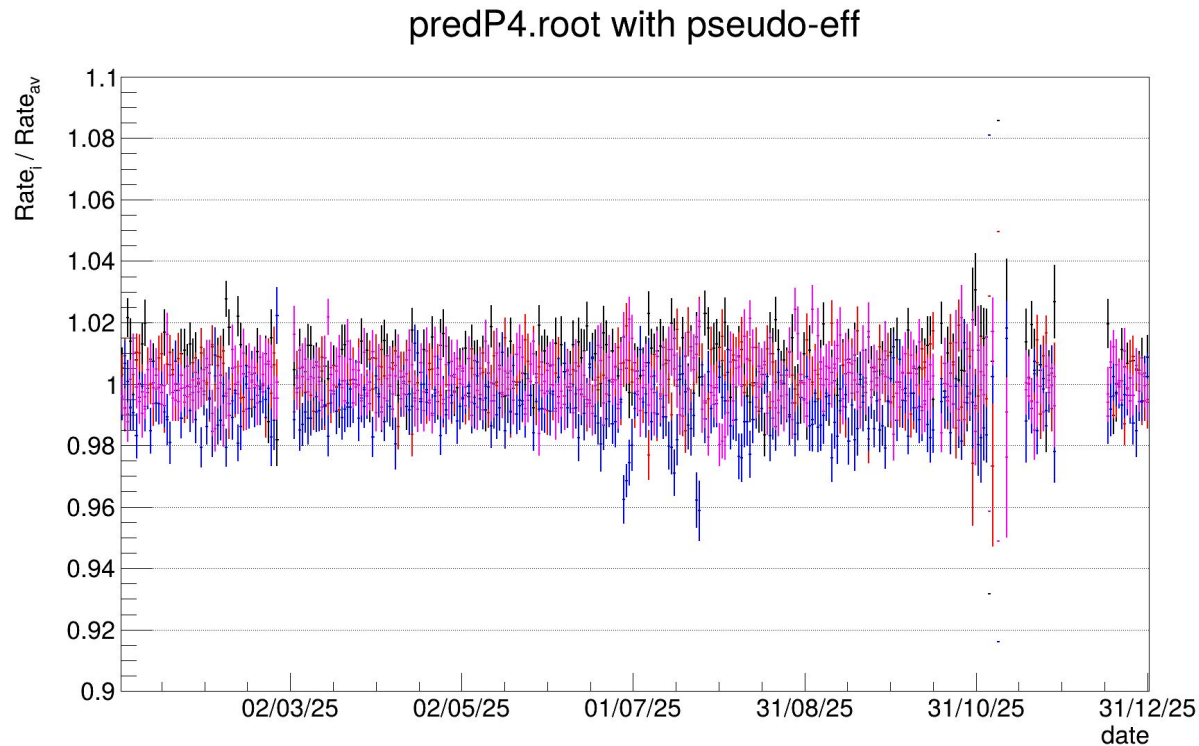
Rate relativo di ogni coppia di piastre verticale
rispetto alla media. Deviazioni da 1 indicano effetti
sistematici dovuti alle efficienze

Grafici di controllo POLA-03 (2025)



Rate relativo di ogni coppia di piastre verticale
rispetto alla media. Deviazioni da 1 indicano effetti
sistematici dovuti alle efficienze

Grafici di controllo POLA-04 (2025)



Rete neurale

Un'idea (lavoro ancora in stato iniziale!) per migliorare la descrizione delle nostre efficienze è quella di rilasciare l'assunzione di indipendenza della risposta dei SiPM sulla stessa piastra.

La proposta è quella di utilizzare

1. Una simulazione veloce per caratterizzare diversi scenari (no correlazioni SiPM, correlazione positiva, correlazione negativa) e ottenere un set di dati come quelli reali: pseudo-efficienze vs rate
2. una rete neurale addestrata sulla simulazione per imparare a riconoscere le correlazioni

Per evitare di addestrare la rete neurale su uno spazio enorme di possibilità la simulazione può utilizzare i valori di pseudo-efficienza misurati per campionare in un intorno della regione che ci interessa.

Simulazione

Una semplice macro di ROOT implementa un MC che simula i rate a partire dalle efficienze dei SiPM ma variando le assunzioni sulla correlazione (positiva o negativa) dei due SiPM sulla stessa piastra:

- Assunzione: rate di muoni con distribuzione $\cos^2(\theta)$ → implica simmetria di rate per combinazioni di piastre simili (extra-info che si potrebbe usare)
- Utilizzata per dare una stima di effetti sistematici con il metodo tradizionale
- Utilizzata per fare training di una rete NN che impari le correlazioni utilizzando i vincoli del sistema (per esempio le simmetrie dei rate) per vedere se riusciamo a misurare le correlazioni e migliorare la stima delle efficienze

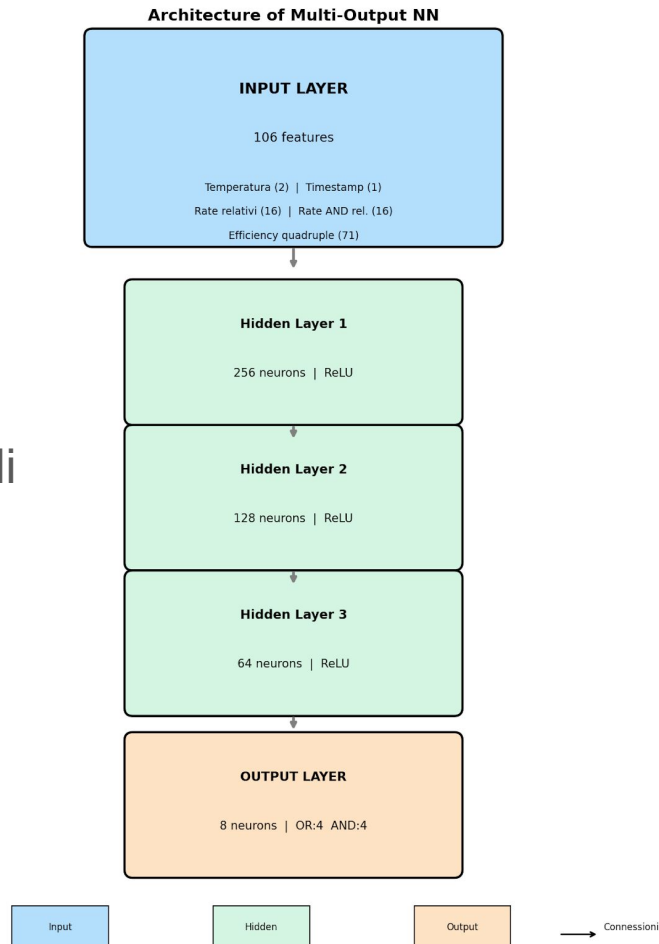
Rete neurale

Architettura (provvisoria, proof of concept):

- 108 input dai dati
- 3 Layer (256, 128, 64)
- 8 output per le 4 combinazioni di piastre verticali: 4 efficienze di Majority, 4 efficienze di AND_4_SiPMs

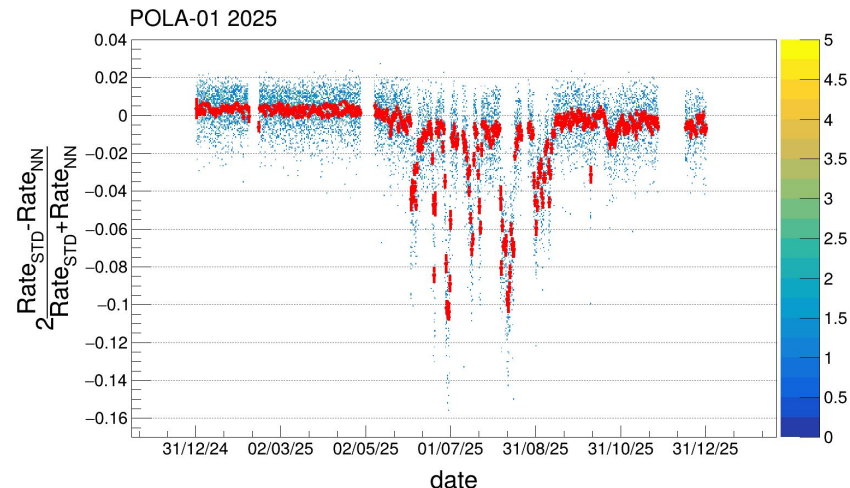
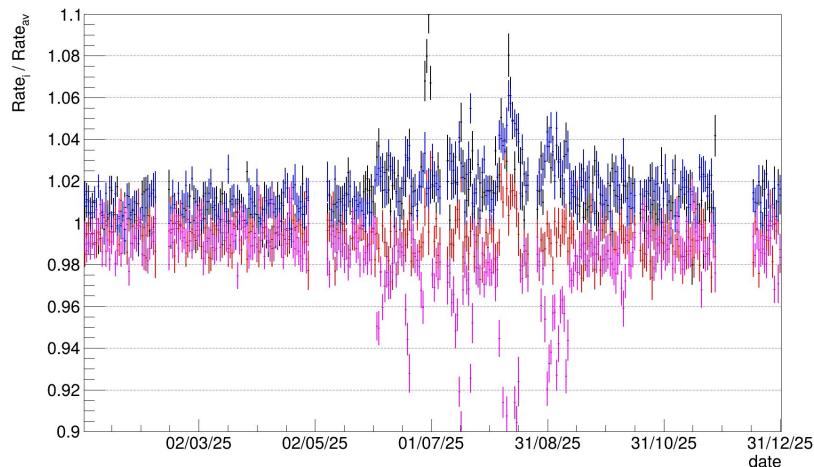
Software:

- MC con ROOT (data driven)
- Macro python per il training (uproot)
- Conversione modello joblib → ROOT
- Inferenza con ROOT sui dati



Primo confronto

predP1.root with pseudo-eff



Il confronto dei rate corretti con il metodo tradizionale e le efficienze delle rete neurale (super-preliminare) mostra che in effetti nella regione estiva (POLA-01 2025) le risposte sono diverse. Ma servono studi più sistematici.

In progress

Per ora ho fatto dei test sul mio laptop ma sto configurando una macchina al CNAF con l'ambiente giusto

- L'ambiente è quasi pronto, rimane da mettere a posto un problema con la conversione del modello NN da python a ROOT per l'inferenza

Una volta validato l'ambiente sulla cloud possono partire studi più sistematici