

Università degli Studi dell'Insubria
Dipartimento di Scienze e Alta Tecnologia
Corso di Laurea Triennale in Fisica

Analisi comparativa di rivelatori Radim 5b

Erica Brisigotti & Ylenia Mascolo



Esame di Laboratorio di Fisica I
Docente: Prof. Dott. Massimo Luigi Maria Caccia
Assistente di Laboratorio: Dott.ssa Samuela Lomazzi
Anno Accademico 2018-2019

Indice

1	Introduzione	3
1.1	Il Radon	3
1.2	Il Radim 5b	4
2	Analisi dati	5
2.1	Verifica della compatibilità dei rilevatori	5
2.1.1	Confronto delle mediane	6
2.1.2	Confronto delle distribuzioni delle medie giornaliere	7
2.2	Determinazione del fattore di scala lineare k	9
2.2.1	<i>Sign test</i> con <i>p-value</i>	9
2.2.2	Sign test con $\tilde{\chi}^2$	11
3	Conclusioni	15
4	Appendice	16
4.1	Media e relativa incertezza	16
4.2	Test del Chi Quadrato	16
4.3	Propagazione degli errori	17
4.4	Pesi	17
5	Fonti	17

1 Introduzione

La seguente esperienza di laboratorio si pone come obiettivo la verifica della compatibilità statistica dei dati acquisiti da cinque dispositivi Radim 5b, posti all'interno di una stessa stanza per rivelare la presenza di radon, un gas altamente pericoloso per l'uomo ma presente di frequente negli ambienti in cui egli vive.

Nonostante i cinque dispositivi siano stati costruiti identicamente e posti nelle medesime condizioni, si osserverà che presentano andamenti simili ma sfalsati a tal punto da risultare incompatibili. Quindi, attraverso diverse strategie di analisi dati, si è determinato un fattore di ricalibrazione k per cui sono riscaldati i dati in modo da renderli statisticamente compatibili.

1.1 Il Radon

Il radon ${}_{86}\text{Rn}$ è un gas nobile naturalmente radioattivo, scoperto da Robert B. Owens e Ernest Rutherford nel 1899, che a temperatura e pressione standard risulta essere inodore, incolore e insapore. Il suo isotopo più stabile (il radon ${}^{222}_{86}\text{Rn}$) fa parte della catena di decadimento α dell'uranio ${}^{238}_{92}\text{U}$: esso è un prodotto del decadimento α del radio ${}^{226}_{88}\text{Ra}$, e può a sua volta decadere α in polonio ${}^{218}_{84}\text{Po}$.

Le concentrazioni di radon sono in media di 10 Bq/m³ in aria, di 50 Bq/m³ in ambienti chiusi e fino a 10⁵ Bq/m³ nel sottosuolo. Siccome l'uranio è presente nelle rocce, nel suolo e nei materiali da costruzione, il radon prodotto dal suo decadimento tende ad accumularsi (in quanto gas molto pesante) raggiungendo concentrazioni potenzialmente pericolose per l'uomo, specialmente in ambienti chiusi o con scarso ricircolo d'aria.

È stata infatti dimostrata l'esistenza di un aumento statisticamente significativo del rischio di carcinoma polmonare correlato all'esposizione prolungata al radon, in ambienti chiusi, per concentrazioni dell'ordine di 100 Bq/m³. Tale rischio risulta ancora maggiore quando l'esposizione al radon è combinata con il consumo di tabacco. Complessivamente, si stima che il radon rappresenti la maggiore fonte di esposizione alle radiazioni ionizzanti e la seconda causa di carcinoma polmonare dopo il tabagismo. Inoltre, nonostante la maggior parte dei tumori ai polmoni dovuti al radon siano causati da concentrazioni medio-basse di radon (siccome risulta più rara l'esposizione ad alti livelli), più alta è la concentrazione di radon nell'ambiente maggiore è il rischio di contrarre carcinomi polmonari e, purtroppo, non esiste una concentrazione minima al di sotto della quale l'esposizione al radon non rappresenti un rischio.

In Italia non sono ancora presenti delle normative nazionali che limitino la concentrazione massima di radon: tutt'al più esiste, a livello nazionale, un livello di attenzione di 150 Bq/m³.

A livello europeo, sono presenti normative che prevedono un livello massimo di 500 Bq/m³ per i luoghi di lavoro mentre raccomandano concentrazioni massime di 200-400 Bq/m³ per edifici residenziali, a seconda che si tratti di nuove realizzazioni o di abitazioni pre-esistenti.

Risulta quindi essenziale il monitoraggio delle concentrazioni di radon nell'aria, che si può ottenere effettuando misure prolungate¹ basate sul suo decadimento α : tale processo è regolato intrinsecamente da una statistica possoniana ed è pertanto legittimo aspettarsi delle fluttuazioni nei conteggi puramente stocastiche, non riconducibili ad errori sperimentali.

¹Per misure prolungate si intendono durate di giorni o settimane, giustificate dal fatto che il tempo di dimezzamento del decadimento ${}^{222}_{86}\text{Rn} \rightarrow \alpha + {}^{218}_{84}\text{Po}$ è pari a circa 4 giorni.

1.2 Il Radim 5b

Le misure che saranno analizzate in questa esperienza sono state acquisite da cinque rivelatori Radim 5b, mostrati in figura (1) e collocati nell'aula di laboratorio di Fisica I dell'Università degli Studi dell'Insubria.

I rivelatori Radim 5b sono misuratori in continuo della concentrazione di radon di sensibilità molto elevata (fino a 20 Bq/m^3) e con tempo di campionamento di 1 ora.

Essi hanno dimensioni $200 \times 150 \times 90 \text{ mm}$ e massa 0.5 kg , e possono acquisire dati per intervalli di tempo molto lunghi (fino ad anni) per via della grande memoria, funzionando a temperature comprese tra 3° e 40° C e con umidità relative dell'aria tra 5% e 99% .

I Radim 5b misurano la concentrazione di radon presente nell'aria tramite dei rivelatori a stato solido, presenti nella camera di rivelazione. Questa è ricoperta di feltro, che filtra l'aria privandola degli altri α -emettitori già presenti al suo interno: il radon, e di conseguenza i suoi prodotti di decadimento (in particolare le particelle α), viene quindi raccolto sulla superficie di un rivelatore a semiconduttore tramite l'applicazione di un campo elettrico ad alta intensità. Il rivelatore a semiconduttore permette infine di calcolare il numero di decadimenti avvenuti a partire dall'intensità misurata.

Siccome nella fase di misura sono stati utilizzati 5 rivelatori differenti, risulta essenziale specificare che i rivelatori necessitano di periodiche ricalibrazioni (in modo da restituire secondo scale uniformi i conteggi misurati), che per esigenze pratiche spesso non risultano simultanee e possono quindi tradursi in alterazioni significative nei dati raccolti dai diversi dispositivi. Per comodità, nell'analisi a seguire si considererà il rivelatore R053 come riferimento in quanto ricalibrato più recentemente rispetto ai restanti R005, R008, R011, R012.

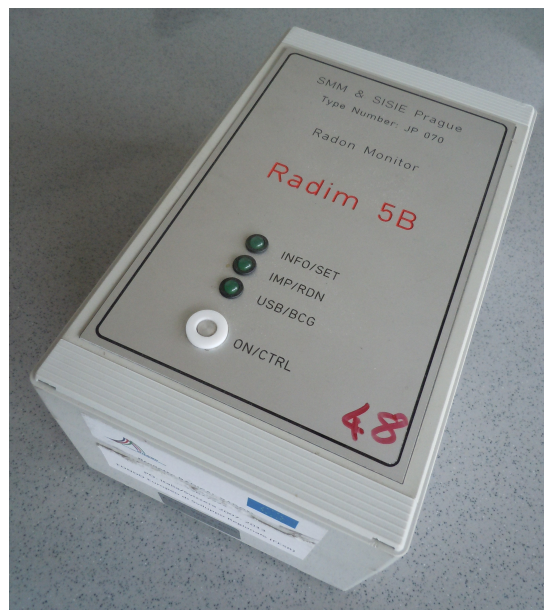


Figura 1: Foto di un rivelatore Radim 5b.

2 Analisi dati

L'analisi dati si è composta di un primo *step* che si concentra sulla verifica della compatibilità dei *set* di dati tramite il confronto delle mediane e delle distribuzioni delle medie giornaliere. Non verificandosi questa condizione, nella fase successiva si individueranno i migliori fattori di ricalibrazione lineare k , ovvero tali che i dati dei quattro rivelatori (R005, R008, R011, R012) risultino compatibili con quelli del quinto (R053, considerato di riferimento e calibrato più recentemente rispetto agli altri).

La solidità di tale modello lineare è stata constatata tramite l'utilizzo di *sign test*, prima coordinato ad un test del p -value e poi in relazione alle proprietà della distribuzione del $\tilde{\chi}^2$, tramite cui è stato possibile ottenere dei valori k_i del coefficiente di calibrazione ottimali.

2.1 Verifica della compatibilità dei rilevatori

La prima fase consiste nel confronto dei diversi *set* di dati $\{x\}_i$ (con $i = 5, 8, 11, 12, 53$) ricavati dai dispositivi Radim 5b, impiegati contemporaneamente nella raccolta dati in zone diverse di una stessa stanza. In particolare, i dati in questione coprono un periodo di circa 800 ore con una frequenza di campionamento pari a circa 1 ora e, trattandosi di conteggi di eventi, questi sono numeri naturali $\{x\}_i \in \mathbb{N}$. Di seguito sono riportati i dati in questione:

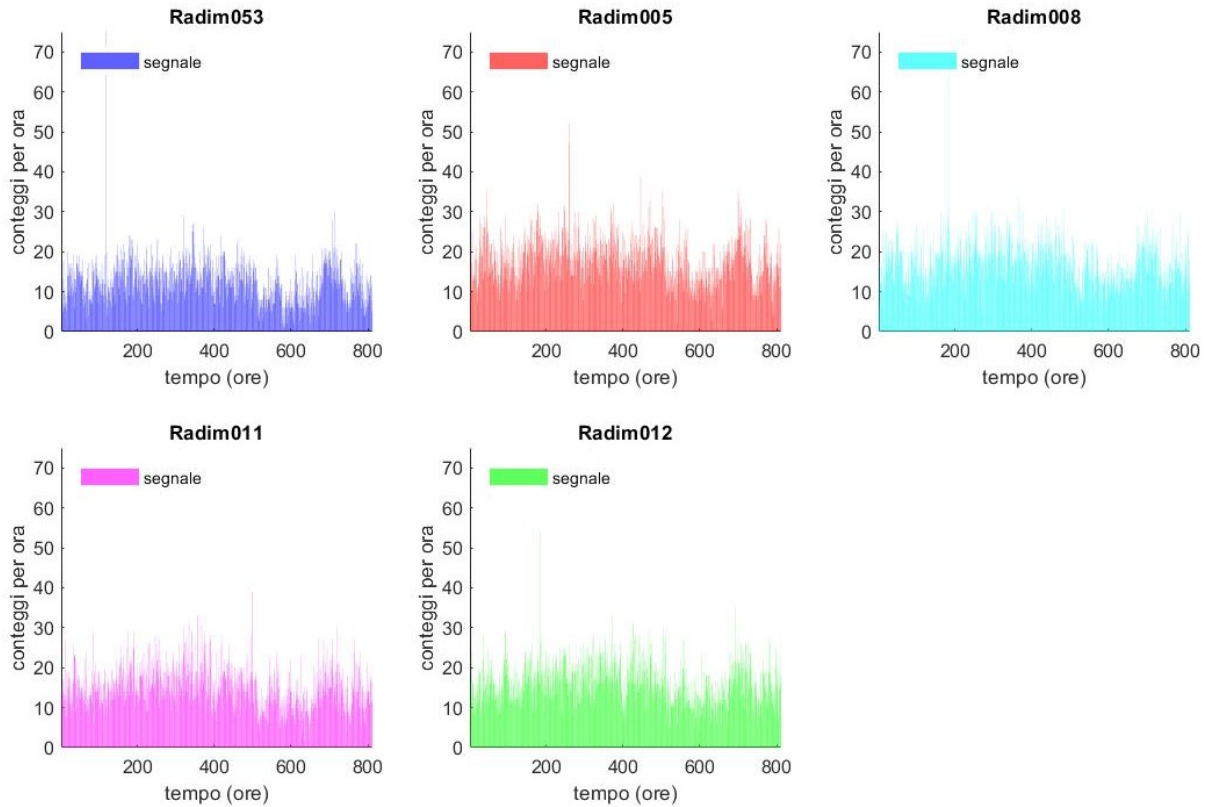


Figura 2: Rappresentazione dei dati $\{x\}_i$ raccolti da ciascun rivelatore ($i = 5, 8, 11, 12, 53$).

Attraverso questa prima analisi sarà possibile stabilire se i dispositivi Radim 5b possono misurare il medesimo andamento orario dei conteggi di eventi (in senso statistico).

2.1.1 Confronto delle mediane

Un primo tentativo di stabilire la compatibilità dei dati raccolti dai dispositivi Radim 5b (in particolare del riferimento $i = 53$ con i restanti), riguarda il confronto dei valori di mediana ottenuti: si tratta infatti del valore centrale della distribuzione del numero di conteggi per ora.

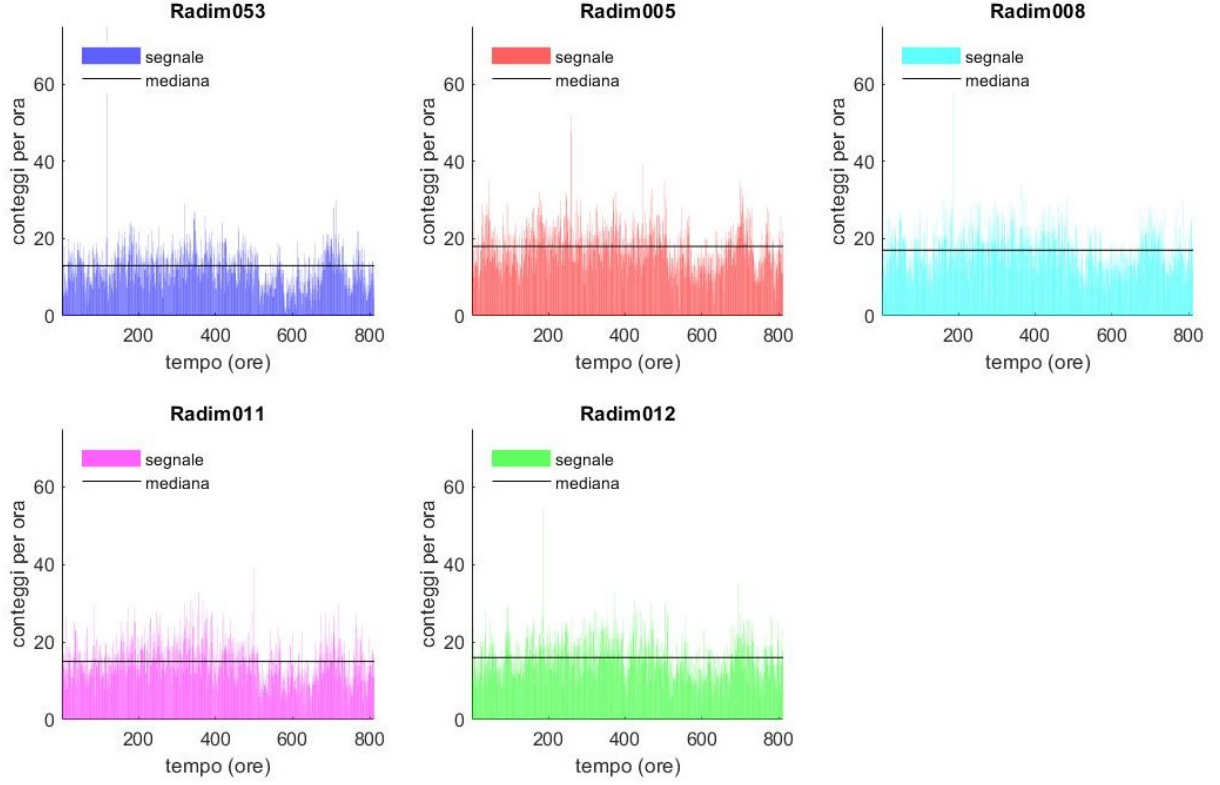


Figura 3: Rappresentazione dei dati $\{x\}_i$ raccolti e della loro mediana, per ciascun rivelatore ($i = 5, 8, 11, 12, 53$).

I valori di mediana x_{med_i} trovati (grazie ad un'apposita funzione di MATLAB) sono stati corretti da errori ricavati come:

$$\sigma_{med_i} = \frac{\text{std}\{x\}_i}{\sqrt{N}} \cdot \frac{\pi N}{2N - 2} \quad (1)$$

dove N è il numero di dati disponibili (ovvero di ore di raccolta dati), mentre $\text{std}\{x\}_i$ corrisponde alla deviazione standard del i -esimo set di dati.

Di seguito si riportano le mediane ottenute, corrette dei rispettivi errori σ_{med} :

$x_{med_{53}}$	x_{med_5}	x_{med_8}	$x_{med_{11}}$	$x_{med_{12}}$
13 ± 0.3	18 ± 0.3	17 ± 0.2	15 ± 0.2	16 ± 0.2

Tabella 1: Valori ottenuti per le mediane x_{med_i} del numero di conteggi per ora, corrette dei rispettivi errori σ_{med} per ciascun rivelatore utilizzato.

Il confronto dei rivelatori ($i = 5, 8, 11, 12$) con il riferimento ($i = 53$) si è quindi basato sul calcolo della distanza in unità di sigma Δ_{σ_i} , trovata come:

$$\Delta_{\sigma_i} = \frac{\Delta_i}{\sigma_i} = \frac{|x_{med_{53}} - x_{med_i}|}{\sqrt{\sigma_{med_{53}}^2 + \sigma_{med_i}^2}} \quad (2)$$

Le distanze per unità di sigma Δ_{σ_i} trovate hanno i seguenti valori:

Δ_{σ_5}	Δ_{σ_8}	$\Delta_{\sigma_{11}}$	$\Delta_{\sigma_{12}}$
12.7	10.4	5.3	7.9

Tabella 2: Valori ottenuti per le distanze in unità di sigma Δ_{σ_i} , per ciascun confronto effettuato.

I valori ottenuti sono significativamente superiori all'usuale soglia di 3 e pertanto è possibile rifiutare (con un livello di confidenza superiore al 99%) la compatibilità dei dati registrati dai quattro rivelatori (R005, R008, R011, R012) con quello di riferimento (R053).

2.1.2 Confronto delle distribuzioni delle medie giornaliere

Il secondo tentativo di stabilire la compatibilità dei dati raccolti dai dispositivi Radim 5b (in particolare del riferimento $i = 53$ con i restanti), consiste nel confronto delle medie (di conteggi per ora) giornaliere: i dati iniziali, relativi ai conteggi per ora corrispondenti a circa 800 ore, sono stati suddivisi in $N_g = 33$ intervalli da 24 ore in ciascuno dei quali è stata effettuata una media, ottenendo i *set* di medie $\{\bar{x}\}_i$ rappresentati di seguito:

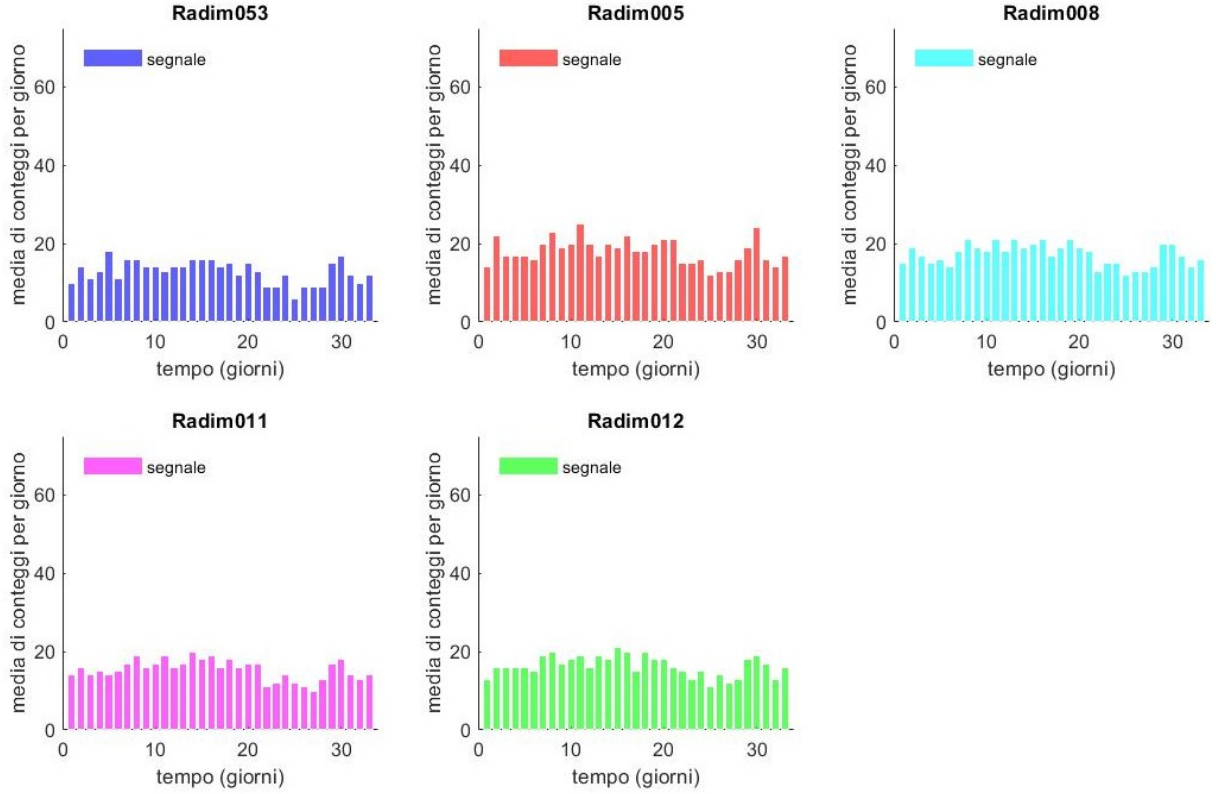


Figura 4: Rappresentazione delle medie giornaliere $\{\bar{x}\}_i$ calcolate, per ciascun rivelatore.

Sono state quindi considerate le distribuzioni dei dati $\{\bar{x}\}_i$, ottenute tramite istogrammi di frequenza: il relativo numero di bins è stato trovato mediante la radice quadrata:

$$n_{bins} = \sqrt{N_g} \sim 5/6 \quad (3)$$

Di seguito sono quindi riportate le distribuzioni delle medie giornaliere, per ciascun dispositivo: alla colonna j -esima (del *set* i -esimo) è stato attribuito un errore $\sigma_{y_{ij}}$ pari a:

$$\sigma_{y_{ij}} = \sqrt{y_{ij}} \quad (4)$$

dove y_{ij} rappresenta l'altezza della colonna dell'istogramma centrata nella posizione z_{ij} .

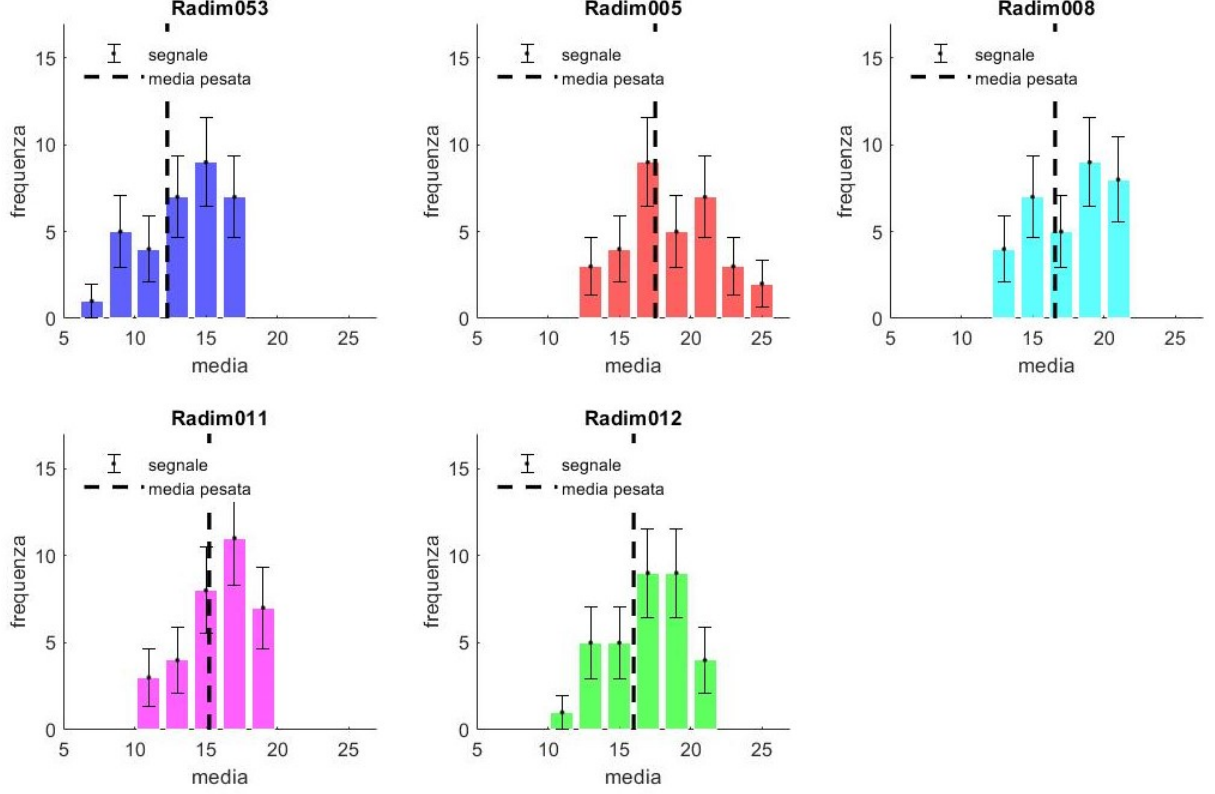


Figura 5: Rappresentazione delle distribuzioni delle medie giornaliere $\{\bar{x}\}_i$ calcolate, per ciascun rivelatore.

I valori medi z_{m_i} delle distribuzioni (rappresentati come rette tratteggiate nel grafico precedente) sono stati individuati applicando la formula della media pesata e corretti da incertezze trovate tramite la propagazione degli errori²:

$$z_{m_i} = \frac{1}{N_g} \sum_{i=1}^{n_{bins}} z_{ij} y_{ij} \quad \sigma_{z_{m_i}} = \frac{1}{N_g} \sqrt{\sum_{i=1}^{n_{bins}} z_{ij}^2 \sigma_{y_{ij}}^2} \quad (5)$$

dove N_g è il numero di intervalli considerati.

Di seguito si riportano i valori medi z_{m_i} ottenuti, corretti dei rispettivi errori $\sigma_{z_{m_i}}$:

$z_{m_{53}}$	z_{m_5}	z_{m_8}	$z_{m_{11}}$	$z_{m_{12}}$
13 ± 0.2	18 ± 0.3	18 ± 0.3	16 ± 0.2	17 ± 0.3

Tabella 3: Valori medi z_{m_i} delle distribuzioni delle medie giornaliere ottenuti per ciascun rivelatore utilizzato, corretti dei rispettivi errori $\sigma_{z_{m_i}}$.

²vedi **4.3 Appendice: Propagazione degli errori.**

Il confronto dei rivelatori ($i = 5, 8, 11, 12$) con il riferimento ($i = 53$) si è quindi basato sul calcolo della distanza in unità di sigma Δ_{σ_i} , trovata come:

$$\Delta_{\sigma_i} = \frac{\Delta_i}{\sigma_i} = \frac{|z_{m_{53}} - z_{m_i}|}{\sqrt{\sigma_{z_{m_{53}}}^2 + \sigma_{z_{m_i}}^2}} \quad (6)$$

Le distanze per unità di sigma Δ_{σ_i} trovate hanno i seguenti valori:

Δ_{σ_5}	Δ_{σ_8}	$\Delta_{\sigma_{11}}$	$\Delta_{\sigma_{12}}$
14.3	12.3	7.9	10.6

Tabella 4: Valori ottenuti per le distanze in unità di sigma Δ_{σ_i} , per ciascun confronto effettuato.

I valori ottenuti sono decisamente superiori all'usuale soglia di 3 e pertanto si rifiuta anche in questo caso (con un livello di confidenza superiore al 99%) la compatibilità dei dati registrati dai quattro rivelatori (R005, R008, R011, R012) con quello di riferimento (R053).

2.2 Determinazione del fattore di scala lineare k

Pur avendo osservato una incompatibilità tra i quattro rivelatori (R005, R008, R011, R012) con quello di riferimento (R053), dall'analisi appena fatta è stato possibile riscontrare un andamento comune a tutti i dispositivi: pertanto si può ipotizzare che stiano misurando il medesimo fenomeno, seppur con calibrazioni differenti.

Il metodo utilizzato per inter-calibrare gli strumenti si è basato sull'introduzione di un coefficiente di ricalibrazione k_i (con $i = 5, 8, 11, 12$), compreso nell'intervallo (0,1) e variato con sensibilità $\delta = 0.0001$ siccome nelle tabelle (1) e (4) i valori di riferimento risultano essere sempre minori dei restanti.

Operativamente il coefficiente di ricalibrazione è stato impiegato come fattore di proporzionalità tra i dati originali $\{x\}_i$ dei rivelatori da calibrare e le quantità $\{h\}_i$ a partire dalle quali sarà possibile trovare una distribuzione da confrontare finalmente con il rivelatore di riferimento.

2.2.1 Sign test con p-value

Il *p-value* corrisponde alla probabilità di ottenere risultati uguali o meno probabili di quelli osservati durante il test, supposta vera l'ipotesi nulla. In altri termini, il valore p aiuta a capire se la differenza tra il risultato osservato e quello ipotizzato è dovuta alla casualità introdotta dal campionamento: se il *p-value* assume valori ridotti (molto prossimi a zero) l'ipotesi nulla perde credibilità, mentre per valori più elevati (comunque compresi nell'intervallo [0,1], siccome si tratta di una probabilità) l'ipotesi nulla è verificata.

Nella pratica questo test è servito per identificare un valore ottimale di fattore di ricalibrazione k_i per ciascun rivelatore da calibrare, in corrispondenza del massimo valore trovato di *p-value*. Infatti, i *p-value* sono stati calcolati³ supponendo come ipotesi nulla l'appartenenza ad una distribuzione del numero di successi, di valore medio μ e deviazione standard σ (ottenuti a partire dalla probabilità $p = 1/2$ e dal numero totale $N = 812$):

$$\mu = Np = 406 \quad \sigma = \sqrt{Np(1-p)} = 14 \quad (7)$$

³utilizzando l'apposita funzione di MATLAB `Z-test(x_{k_i}, μ, σ)`.

Il valore confrontato con questa distribuzione corrisponde a x_{k_i} : esso equivale al numero di volte in cui i dati riscaldati $k_i \cdot \{x\}_i$ (ottenuti da quelli registrati dal rivelatore da calibrare) sono risultati maggiori dei corrispondenti dati $\{x\}_{53}$ del rivelatore di riferimento, fissato un valore k_i .

Di seguito sono rappresentati i p -value in funzione dei corrispondenti fattori di scala k_i :

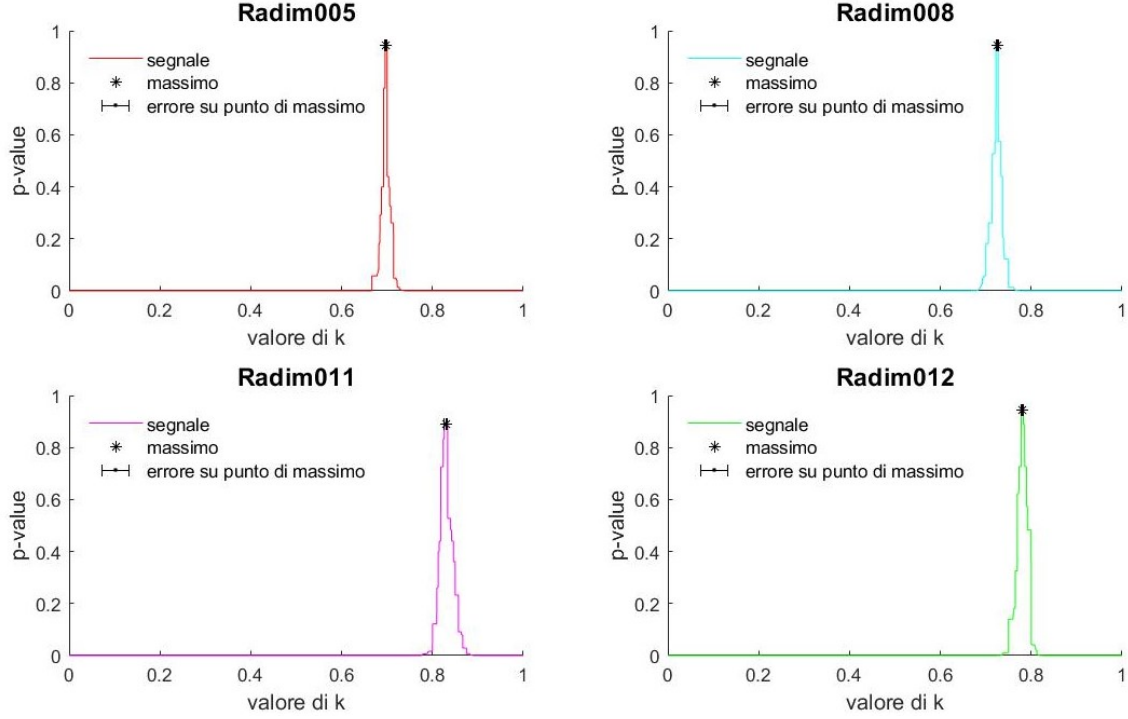


Figura 6: Rappresentazione dei valori di p -value ottenuti per ciascun valore di k_i considerato, con identificazione del punto di massimo (cioè del valore ottimale di k_i), per ciascun rivelatore.

Come è possibile osservare dalla figura precedente, i valori ottimali di coefficiente di ricalibrazione k_i sono stati dotati di un errore pari alla metà della larghezza dei *plateau* corrispondente, dovuti al fatto che i dati di partenza sono numeri naturali. Di seguito sono riportati tali valori e i corrispondenti p -value, che si osservano essere prossimi a 1:

	R005	R008	R011	R012
k ottimale	0.698 ± 0.002	0.725 ± 0.003	0.830 ± 0.004	0.780 ± 0.002
p -value massimo	0.9440	0.9440	0.8884	0.9440

Tabella 5: Valori ottenuti di fattore di scala k ottimale e di p -value massimo per ciascun rivelatore.

Per avere un'ulteriore conferma che i valori di fattore di scala ottimali k_i siano sensati, è stato analizzato il numero di successi x_{k_i} in funzione dei possibili valori di fattore di scala, come riportato in figura (7): infatti x_{k_i} è definito come il numero di volte in cui i dati riscaldati $k_i \cdot \{x\}_i$ sono risultati maggiori dei corrispondenti dati $\{x\}_{53}$ del rivelatore di riferimento.

Tali andamenti sono stati confrontati con il valore medio $\mu = 406$ di una distribuzione binomiale di probabilità $p = 0.5$ e numero totale $N = 812$. In generale, infatti supponendo che due rivelatori siano inter-calibrati questi avranno uguale probabilità che la singola misura registrata dal primo sia maggiore di quella registrata dal secondo, e viceversa: pertanto contando quante volte si verifica una delle due evenienze ci si aspetta un valore medio pari a $\mu = Np = 406$.

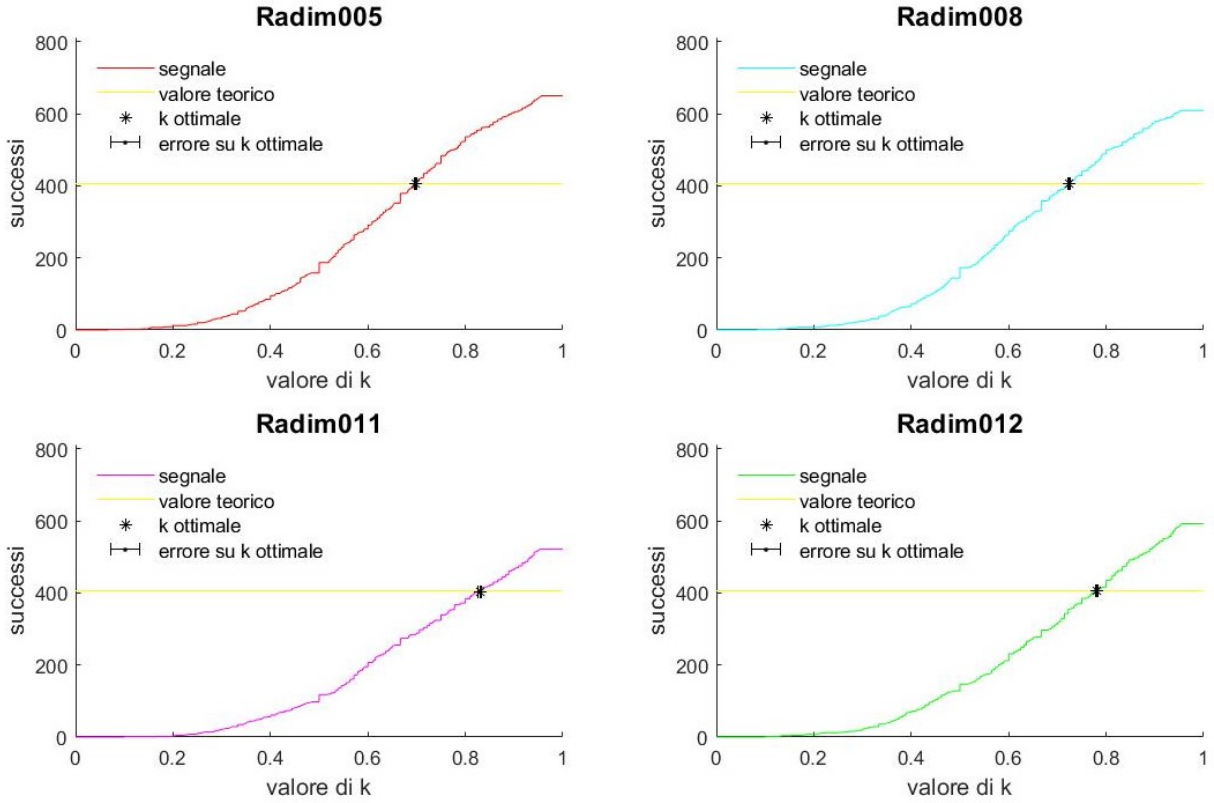


Figura 7: Rappresentazione del numero di successi al variare dei possibili fattori di scala, con evidenziato il numero di successi corrispondente al miglior coefficiente k_i (riportati in tabella (5)), per ciascun rivelatore.

Come è possibile osservare dalla figura (7), i valori di x_{k_i} corrispondenti ai fattori k_i ottimali risultano visibilmente prossimi al valore di aspettazione $\mu = 406$. Per completezza si riportano di seguito i valori di x_{k_i} appena menzionati:

x_{k_5}	x_{k_8}	$x_{k_{11}}$	$x_{k_{12}}$
405	407	404	407

Tabella 6: Numero di successi corrispondente al miglior coefficiente k_i (riportati in tabella (5)), per ciascun rivelatore.

2.2.2 Sign test con $\tilde{\chi}^2$

Un metodo alternativo per ricavare il fattore di scala k_i ottimale prevede l'utilizzo del test del chi-quadrato ridotto $\tilde{\chi}^2$. In generale, a partire da un campione di valori misurati e da una qualunque ipotesi teorica, è possibile applicare la statistica del $\tilde{\chi}^2$ come test di adattamento, cioè per confrontare la distribuzione osservata su un campione e la distribuzione teorica che si ipotizza descrivere la popolazione da cui proviene il campione.

Per questa analisi si considera come ipotesi teorica la compatibilità di ciascun rivelatore con quello di riferimento: nella pratica corrisponde alla compatibilità tra la distribuzione binomiale⁴ e la distribuzione del numero di successi, partizionati in 80 intervalli da 10 ore.

Il numero di successi corrispondenti a ciascun intervallo è stato trovato come il numero di volte n (con $0 \leq n \leq 10$) in cui i dati riscaldati $k_i \cdot \{x\}_i$ relativi al rivelatore i -esimo sono risultati maggiori dei corrispondenti dati $\{x\}_{53}$ del rivelatore di riferimento.

Di seguito si rappresentano le distribuzioni del numero di successi iniziali, ottenute in assenza di riscaldamento ($k_i=1$), per ciascun rivelatore in questione:

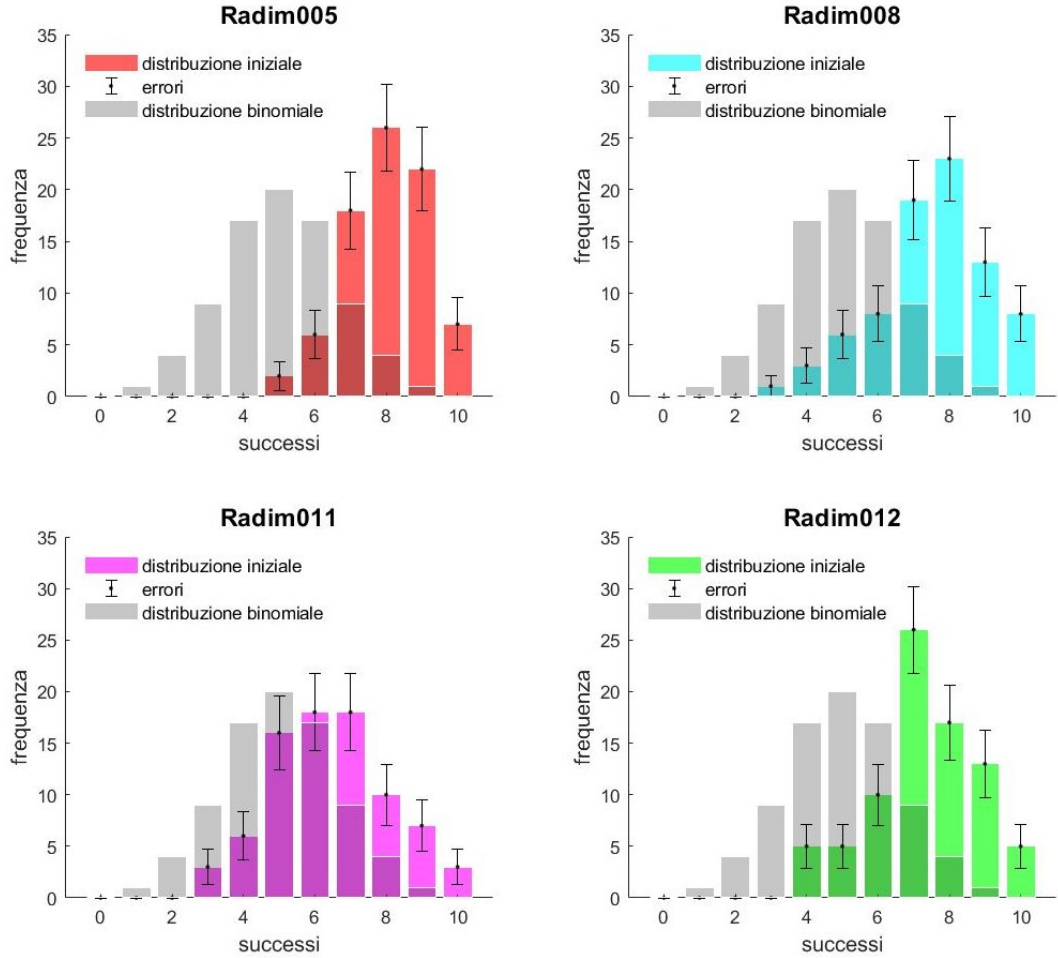


Figura 8: Rappresentazione delle distribuzioni di successi iniziali in assenza di riscaldamento e della distribuzione teorica binomiale, per ciascun rivelatore ($i = 5, 8, 11, 12$)

Dalle distribuzioni in figura (8) è possibile osservare che i valori più elevati di mediane e medie delle medie (ricavati nella fase di confronto e riportati nelle tabelle (1) e (3)) si traducono anche in uno spostamento della distribuzione dei successi verso numeri di successi più alti.

Sono quindi giustificati anche in questo caso dei valori di k_i compresi nell'intervallo $(0,1)$, in modo da arretrare la distribuzione dei successi allineandola il più possibile a quella teorica.

⁴Per distribuzione binomiale di successi si intende $80 \cdot \mathbb{B}(n, N; p)$, dove $\mathbb{B}(n, N; p)$ è la distribuzione di probabilità binomiale di probabilità $p = 0.5$ e numero di eventi possibili $N = 10$ mentre 80 è il numero di ripetizioni effettuate.

Il valore di k_i ottimale è stato poi individuato in corrispondenza del minimo valore di $\tilde{\chi}^2$ ottenuto: la statistica di $\tilde{\chi}^2$, infatti, prevede un valore tanto minore quanto più dati teorici e sperimentali risultano prossimi tra loro, ovvero se la teoria spiega sufficientemente bene il campione.

Di seguito si riportano gli andamenti del $\tilde{\chi}^2$ ottenuti per ciascun valori di k_i considerato:

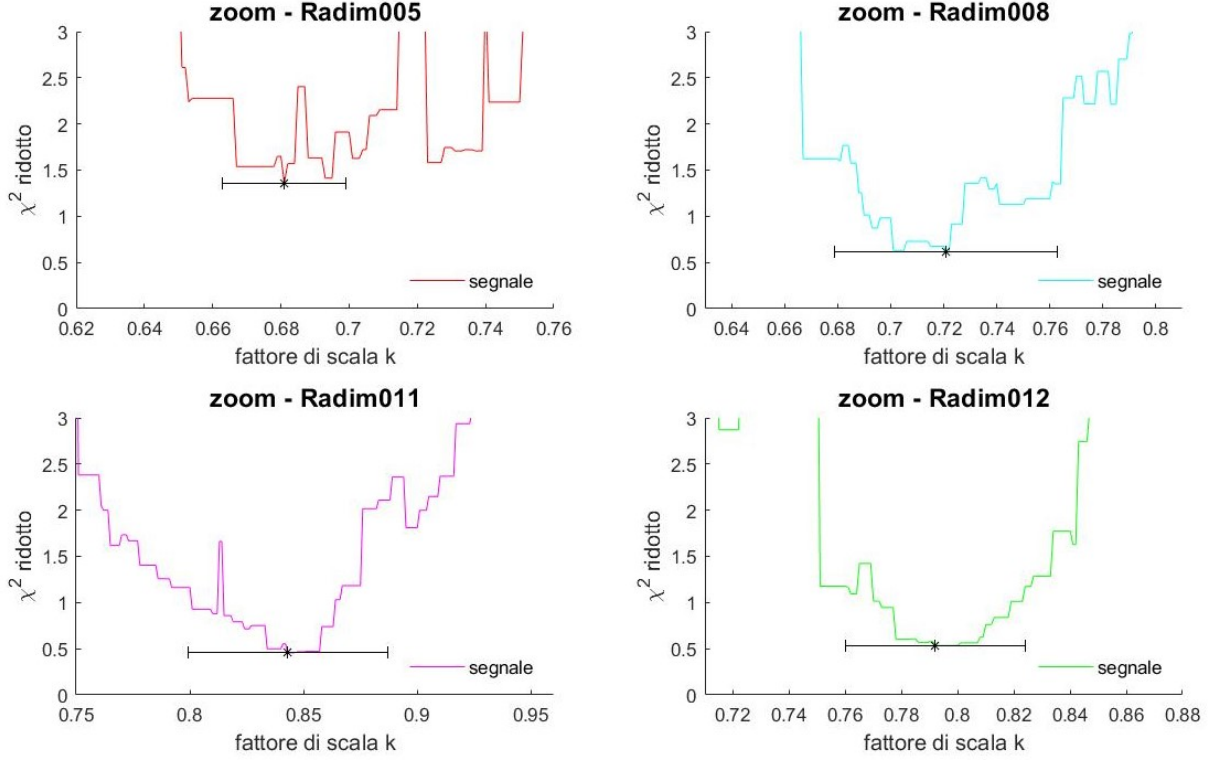


Figura 9: Rappresentazione dell'andamento di $\tilde{\chi}^2$ in funzione del fattore di scala k_i , con evidenziato il valore ottimale e il relativo errore, per ciascun rivelatore da calibrare.

Come è possibile osservare nella figura precedente, in alcune porzioni di grafico è presente un comportamento a gradoni dovuto alla quantizzazione del numero di successi ($n \in \mathbb{N}$).

Inoltre, è stato possibile trovare un unico valore ottimale di fattore di scala k_i per ciascun rivelatore da calibrare: questi sono stati riportati in tabella (7) assieme alle corrispettive incertezze, calcolate come gli intervalli corrispondenti ad incrementi unitari del valore di $\tilde{\chi}^2$ (rispetto al suo valore minimo).

	R005	R008	R011	R012
k ottimale	0.68 ± 0.02	0.72 ± 0.04	0.84 ± 0.04	0.79 ± 0.03
p -value massimo	1.3615	0.6143	0.4551	0.5345

Tabella 7: Valori ottenuti di fattore di scala k ottimale e di minimo $\tilde{\chi}^2$ per ciascun rivelatore da calibrare con il riferimento.

I minimi valori di $\tilde{\chi}^2$ ottenuti risultano prossimi o minori del valore di aspettazione di 1.

Si deduce quindi che, per ciascun rivelatore da calibrare, le distribuzioni ottenute riscaldando i dati iniziali $\{x\}_i$ per il fattore di scala ottimale k_i (rappresentate in figura (10)) risultano compatibili con la distribuzione binomiale (data da $80 \cdot \mathbb{B}(n, N = 10; p = 0.5)$): lo stesso vale quindi per i dati riscaldati $k_i \cdot \{x\}_i$ da cui esse sono state ottenute, che possono essere considerati compatibili con i dati $\{x\}_{53}$ acquisiti dal rivelatore di riferimento R053.

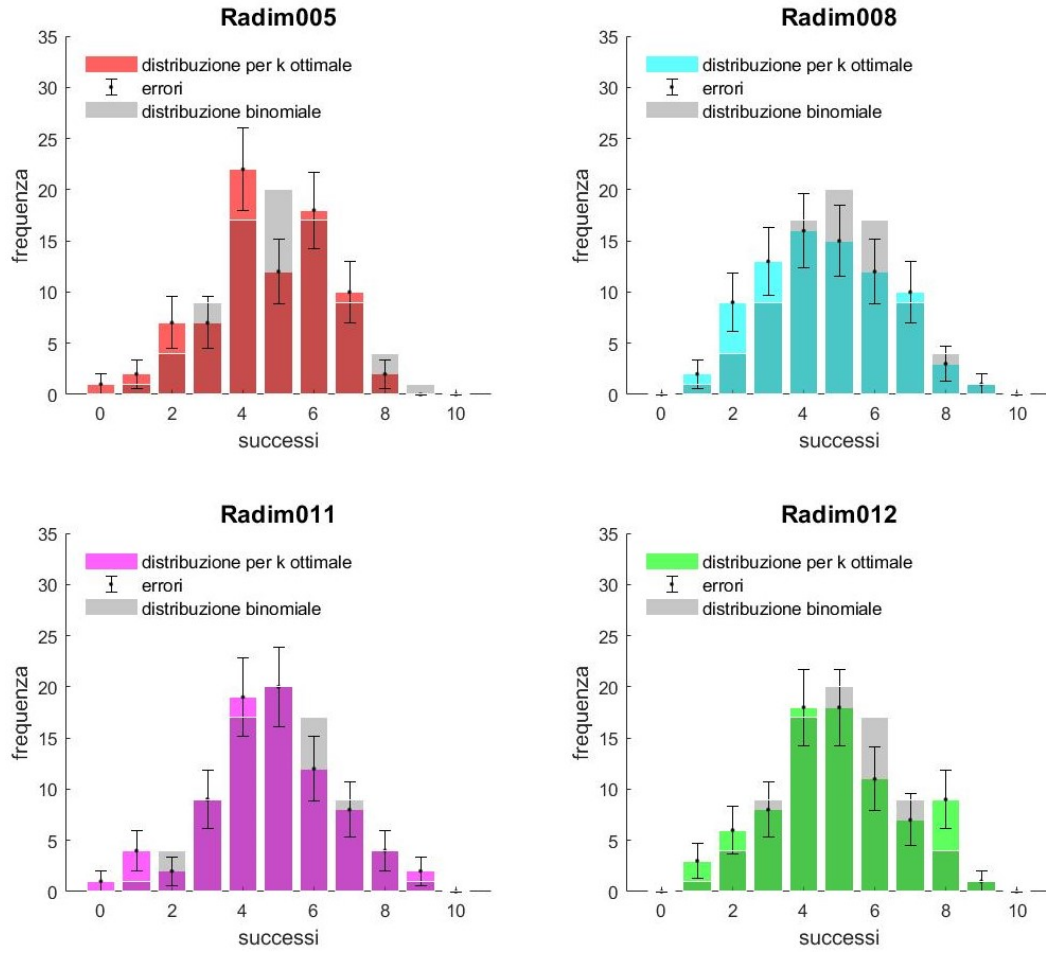


Figura 10: Rappresentazione delle distribuzioni di successi ottenute riscalando i dati iniziali $\{x\}_i$ per il fattore di scala ottimale k_i e della distribuzione teorica binomiale, per ciascun rivelatore da ricalibrare ($i = 5, 8, 11, 12$).

3 Conclusioni

L'esperienza di laboratorio appena illustrata si è rivelata soddisfacente in quanto l'incompatibilità statistica osservata nella prima fase di analisi è stata superata tramite l'introduzione di un fattore di ricalibrazione k_i per ciascun rivelatore da calibrare con il riferimento R053.

Tale approccio, giustificato dall'osservazione di un andamento analogo dei dati, si è rivelato esauriente siccome è stato possibile trovare due valori ottimali di fattore di scala k_i (uno per ciascun metodo utilizzato): questi hanno permesso di affermare che dati presi dai vari dispositivi risultano compatibili con quelli misurati dal rivelatore di riferimento, ovvero che hanno registrato fenomeni analoghi. Pertanto si è dedotto che la distribuzione di radon all'interno della stanza sia stata uniforme, anche perchè in caso contrario non sarebbe stato possibile individuare in modo preciso i fattori di ricalibrazione ottimali k_i (a cui sarebbero corrisposte incertezze eccessivamente elevate e valori di p -value e $\tilde{\chi}^2$ inverosimili).

Infine, per fornire un unico valore di fattore di ricalibrazione k_i per ciascun rivelatore ($i = 5, 8, 11, 12$), si discute a seguire la compatibilità dei valori ottenuti in ciascun metodo, riportati nelle tabelle (5) e (7). In particolare, sono stati calcolati la distanza $\Delta_i = |k_i^p - k_i^{\chi^2}|$ presente tra i valori trovati nei due metodi e l'errore combinato σ_i :

$$\sigma_i = \sqrt{(\sigma_i^p)^2 + (\sigma_i^{\chi^2})^2} \quad (8)$$

in modo tale da ottenere gli scarti normalizzati Δ_{σ_i} , calcolati come:

$$\Delta_{\sigma_i} = \frac{\Delta_i}{\sigma_i} \quad (9)$$

e riportati nella tabella a seguire:

	R005	R008	R011	R012
k^p	0.698 ± 0.002	0.725 ± 0.003	0.830 ± 0.004	0.780 ± 0.002
k^{χ^2}	0.68 ± 0.02	0.72 ± 0.04	0.84 ± 0.04	0.79 ± 0.03
Δ	0.018	0.005	0.01	0.01
σ	0.02	0.04	0.04	0.03
Δ_{σ}	0.90	0.13	0.25	0.33

Tabella 8: Valori ottenuti per i fattore di scala k_i ottimali in ciascun metodi, utilizzati per calcolare gli scarti normalizzati Δ_{σ_i} a partire dalle distanze Δ_i e dagli errori combinati σ_i per ciascun rivelatore da calibrare con il riferimento.

Siccome le distanze in unità di sigma Δ_{σ_i} trovate risultano molto inferiori rispetto all'usuale soglia di 3, è possibile stabilire la compatibilità tra i fattori di scala ottimali k_i trovati tramite le due strategie. Pertanto è possibile calcolare una media dei due valori, riportata in tabella:

	R005	R008	R011	R012
\bar{k}	0.69 ± 0.02	0.72 ± 0.04	0.84 ± 0.04	0.79 ± 0.03

Tabella 9: Valori medi calcolati per il fattore di scala ottimale k_i , necessario per ricalibrare il rivelatore in questione con il riferimento R053.

4 Appendice

4.1 Media e relativa incertezza

Sia x_i un *set* di N dati. Si definisce *valore medio*:

$$\bar{x} = \frac{1}{N} \cdot \sum_{i=1}^N x_i \quad (10)$$

Inoltre, è possibile attribuire a questo valore un'incertezza pari a:

$$\sigma_{\bar{x}} = \frac{\text{std}(x_i)}{\sqrt{N}} \quad (11)$$

Dove $\text{std}(x_i)$ è la *deviazione standard* che può essere calcolata tramite la seguente formula:

$$\text{std}(x_i) = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^N |x_i - \bar{x}|^2} \quad (12)$$

4.2 Test del Chi Quadrato

Il *Test del Chi Quadrato* è un test di verifica delle ipotesi utilizzato in statistica, basato sulla distribuzione del χ^2 .

Siano un *set* di dati sperimentali $\{x^s\}$, caratterizzati da incertezze $\{\sigma^s\}$, e dei valori attesi $\{x^t\}$; si definisce il *Chi Quadrato* come:

$$\chi^2 = \sum_{i=1}^n \omega_i^s \cdot (x_i^s - x_i^t)^2 \quad (13)$$

dove i pesi usati nell'equazione precedente possano essere calcolati, a partire dagli errori $\{\sigma^s\}$, come:

$$\omega_i^s = \frac{1}{(\sigma_i^s)^2} \quad (14)$$

Il valore utile per la verifica delle ipotesi è però il *Chi Quadrato Ridotto* che viene calcolato come:

$$\tilde{\chi}^2 = \frac{\chi^2}{\text{dfe}} \quad (15)$$

dove il numero di *gradi di libertà* dfe si ottiene a partire dal numero di dati considerati n_{dati} , dal numero di parametri del *fit* n_{par} e dal numero di punti esclusi n_{esc} ; la formula, considerati tali parametri, risulta essere:

$$\text{dfe} = n_{\text{dati}} - n_{\text{par}} - n_{\text{esc}} \quad (16)$$

In generale, il valore di aspettazione per questa distribuzione è di 1, pertanto per $\tilde{\chi}^2 \gg 1$ si può dedurre che la distribuzione sperimentale non segua quella teorica ipotizzata.

4.3 Propagazione degli errori

Sia una funzione continua $f(x_1, \dots, x_n)$ dipendente da n variabili x_1, \dots, x_n , caratterizzate rispettivamente da incertezze $\sigma_1, \dots, \sigma_n$. Supponendo che tali incertezze possano essere ritenute indipendenti e casuali, è possibile stimare l'incertezza σ_f sul valore di $f(x_1, \dots, x_n)$ nel seguente modo:

$$\sigma_f = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)^2 \cdot \sigma_i^2} \quad (17)$$

4.4 Pesi

Data una serie di dati $\{y_i\}$ e il relativo *set* di incertezze $\{\sigma_i\}$, è possibile ricavare i pesi $\{w_i\}$ da attribuire a $\{y_i\}$ come:

$$w_i = \begin{cases} \frac{1}{\sigma_i^2} & \sigma_i \neq 0 \\ 0 & \sigma_i = 0 \end{cases} \quad (18)$$

5 Fonti

1. John R. Taylor, *Introduzione all'analisi degli errori*, 1999
2. Maria Garetto, *Statistica: lezioni ed esercizi*, 2002
3. Marco Mombelli, *Radiazioni ionizzanti 9: Radon*