

## **CAMPAGNA DI MISURA 2009 -2010 DELLA CONCENTRAZIONE DI RADON INDOOR IN LOMBARDIA**

A. Alberici<sup>1</sup>, S. Arrigoni<sup>2</sup>, R. Borgoni<sup>3</sup>, A. Cugini<sup>2</sup>, M. Di Toro<sup>2</sup>, D. de Bartolo<sup>1</sup>, M. Romanelli<sup>2</sup>, V. Tritto<sup>3</sup>

<sup>1</sup> ARPA della Lombardia, Settore Aria e Agenti Fisici, Milano

<sup>2</sup> ARPA della Lombardia, Dipartimento provinciale di Bergamo

<sup>3</sup> Università degli Studi di Milano – Bicocca, Dipartimento di Statistica

### **ABSTRACT**

In questo lavoro vengono presentati i risultati della nuova campagna regionale di misura della concentrazione di radon indoor realizzata nel 2009/2010 in Lombardia. Tale monitoraggio è una delle attività che la Regione Lombardia ha affidato ad ARPA nell'ambito dell'avvio del Piano Nazionale Radon per la riduzione del rischio di tumore polmonare in Italia.

Questa nuova campagna è stata condotta al fine di migliorare e approfondire le conoscenze sulla distribuzione territoriale del radon indoor negli edifici lombardi. Nello specifico, sono stati definiti cinque obiettivi, tra i quali la valutazione della rappresentatività della precedente campagna di monitoraggio condotta nel biennio 2003/2004, il confronto tra le concentrazioni di radon in locali a diversi piani di un edificio, il miglioramento delle informazioni spaziali sul territorio, e un confronto, in merito all'analisi dei dati, tra l'approccio geostatistico recentemente introdotto in questo ambito, con quello statistico tradizionale.

Per ognuno degli obiettivi è stato definito un piano di campionamento specifico e, in totale, sono state misurate circa 1000 unità immobiliari dislocate in 7 province della regione (Bergamo, Brescia, Lodi, Mantova, Milano, Sondrio e Varese). In particolare, in questo lavoro vengono presentati i risultati relativi a due degli obiettivi prefissati, quale la valutazione della rappresentatività dei dati della precedente campagna e il confronto delle concentrazioni ai diversi piani di un edificio.

### **OBIETTIVI DELLA CAMPAGNA**

Nell'ambito delle attività della Regione Lombardia connesse con l'avvio del Piano Nazionale Radon per la riduzione del rischio di tumore polmonare in Italia (PNR-CCM, piano che il CCM, Centro Nazionale per il controllo e la prevenzione delle Malattie, ha affidato all'Istituto Superiore di Sanità), è emersa la necessità di approfondire ulteriormente le conoscenze sulla Lombardia, allo scopo di avere informazioni più precise sulla distribuzione territoriale della concentrazione di radon indoor. Si è voluto quindi approfondire la campagna di misura regionale di radon indoor condotta nel 2003/2004, con una nuova campagna realizzata nel 2009/2010, il cui piano di campionamento è stato discusso e concordato con la DG Sanità della Regione Lombardia e con l'Istituto Superiore di Sanità (ISS).

Nello specifico, questa nuova campagna è stata condotta per rispondere a cinque obiettivi:

- 1) valutare la rappresentatività dei dati della campagna 2003/2004 con riferimento ad unità immobiliari site al piano terra: la campagna 2003/2004, infatti, aveva preso in considerazione solo questa tipologia di locali e la scelta dei siti di misura era avvenuta in modo non casuale;
- 2) ottenere una stima rappresentativa, rispetto alla popolazione residente, della concentrazione di radon indoor sul territorio;
- 3) stimare i rapporti tra le concentrazioni di radon in abitazioni ai diversi piani di un edificio;
- 4) ottenere ulteriori informazioni spaziali sul territorio tramite un approccio geostatistico;
- 5) confrontare localmente l'approccio geostatistico con tecniche statistiche convenzionali.

Per ognuno di questi obiettivi è stato predisposto un piano di campionamento specifico. In particolare, i punti di misura sono stati distinti in 4 tipologie, a1, a2, b e c, a seconda del criterio di campionamento e dell'obiettivo perseguito:

- Tipologia a1 (Obiettivo 1 – Obiettivo 2):

scelta di alcuni comuni lombardi tra quelli in cui fossero state effettuate un certo numero di misure (possibilmente almeno 10) durante la campagna 2003/2004;

estrazione casuale di famiglie dall'elenco dell'anagrafe nei comuni scelti, senza vincoli sul piano dell'unità immobiliare della famiglia estratta;

- Tipologia a2 (Obiettivo 3): qualora l'edificio della famiglia estratta (a1) fosse costituito da più piani, dovevano essere posizionati rivelatori anche nelle altre unità immobiliari, possibilmente della stessa colonna, fino al 3° piano e uno anche all'ultimo piano. La stessa procedura è stata prevista anche per le case monofamiliari disposte su più piani;

- Tipologia b (Obiettivo 4): a partire da coppie di coordinate opportunamente scelte al fine di migliorare la stima della concentrazione nelle aree territoriali e, quindi, la costruzione di mappe tematiche, sono stati identificati gli indirizzi (più di uno, per avere alternative in caso di rifiuto) più vicini alle coordinate stesse. I rivelatori dovevano, quindi, essere posizionati in unità immobiliari al piano terra fra gli indirizzi indicati;

- Tipologia c (Obiettivo 5): estrazione casuale delle famiglie in alcuni comuni fra quelli in cui ricadono punti di misura di tipologia b) e scelta di unità immobiliari site al piano terra.

I comuni della Lombardia sono stati convenzionalmente suddivisi in 5 diverse categorie, sulla base dei valori di concentrazione di radon assegnati in seguito ai risultati della campagna 2003/2004 (Borgoni *et al.*, 2007). Nello specifico, un comune appartiene alla:

- *fascia alta* se il valore medio di concentrazione di radon indoor stimato è maggiore di 183 Bq/m<sup>3</sup> (80° percentile della distribuzione territoriale delle concentrazioni medie comunali);
- *fascia medio-alta* se il valore medio è compreso fra 145 e 183 Bq/m<sup>3</sup> (60° e 80° percentile);
- *fascia media* se il valore medio è compreso fra 129,6 e 145 Bq/m<sup>3</sup> (40° e 60° percentile);
- *fascia medio-bassa* se il valore medio è compreso fra 109 e 129,6 Bq/m<sup>3</sup> (20° e 40° percentile);
- *fascia bassa* se il valore medio è minore di 109 Bq/m<sup>3</sup> (20° percentile).

Il campionamento, per questa nuova campagna, è avvenuto in 77 comuni scelti all'interno delle fasce alta, media e bassa e localizzati nelle provincie di Bergamo, Brescia, Lodi, Mantova, Milano, Sondrio e Varese.

I punti di misura dovevano essere unità immobiliari costituite da abitazioni o luoghi di lavoro assimilabili ad abitazioni per caratteristiche e per utilizzo (es.: uffici). Per motivi di modalità di utilizzo e aerazione, non sono state considerate idonee: case di vacanza, negozi, garage, rimesse, né gli atri di accesso alle strutture pubbliche e, in generale, si escludevano locali aventi volumetrie superiori a 300 m<sup>3</sup>. Analogamente all'interno delle unità immobiliari, i rilevatori non dovevano essere posizionati in ripostigli o locali similari, con scarsa aerazione e frequentazione. Locali come bagni e cucine sono stati esclusi dall'indagine sia a causa delle particolari condizioni di utilizzo e di aerazione, sia a causa dell'elevata umidità che potrebbe, in casi estremi, influenzare i risultati delle misure. Qualora si trattasse di abitazione, si doveva scegliere come locale di misura preferenzialmente la camera da letto.

Le misure, effettuate tramite i rilevatori a tracce di tipo CR-39, hanno avuto durata annuale (con cambio semestrale del dosimetro allo scopo di evitare un'eventuale saturazione del rivelatore). In alcuni punti (circa l'8%) è stato posizionato un dosimetro aggiuntivo a scopo di controllo.

I dosimetri sono stati distribuiti alle ASL a settembre 2009 e il posizionamento è avvenuto, per la maggior parte dei casi, fra la metà di settembre e la fine di ottobre. Il cambio semestrale dei dosimetri, dunque, è perlopiù avvenuto fra marzo e aprile 2010.

La lettura dei dosimetri è stata effettuata dal laboratorio radiometrico del Dipartimento ARPA di Bergamo e i dati finali, comprendenti le misure di entrambi i semestri, sono stati resi disponibili verso metà marzo 2011. Il Dipartimento di Statistica dell'Università degli Studi di Milano-Bicocca si è occupato delle analisi statistiche e geostatistiche.

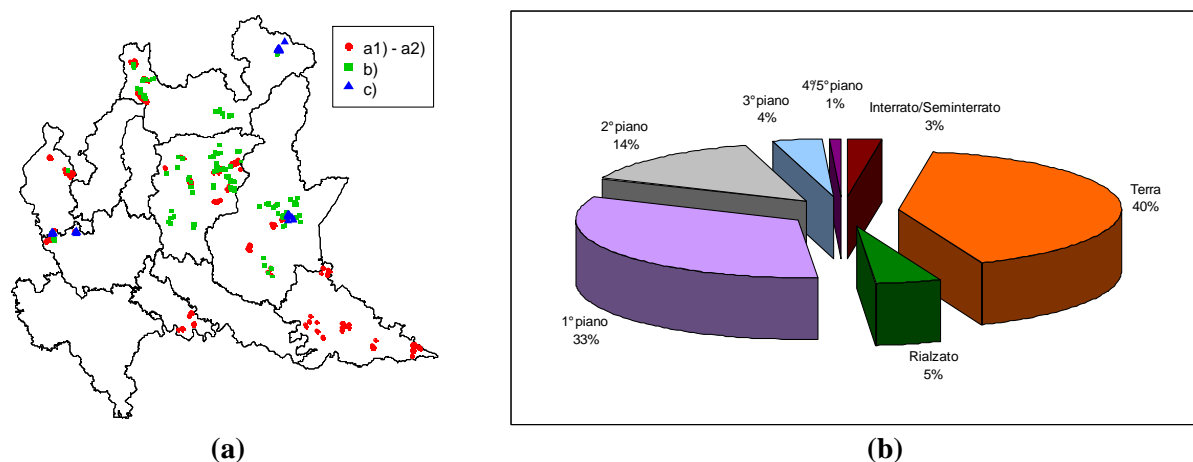
## I DATI

Il database, contenente tutte le informazioni circa le caratteristiche dell'edificio, del locale di misura e dei valori di concentrazione di radon indoor semestrali, è costituito da 1040 record. Sono stati considerati validi a fini dell'analisi 966 punti di misura, distribuiti nel territorio come mostrato in Fig. 1-a, dove viene anche riportata la loro ripartizione per tipologia di campionamento.

La Tab. 1 riporta il numero di punti effettuati per provincia e tipologia di campionamento, mentre la Tab. 2 la loro ripartizione per fascia. Come si evince da quest'ultima tabella, nella nuova campagna sono state campionate maggiormente unità immobiliari in comuni di fascia alta (73%) rispetto alla campagna del 2003/2004 in cui i punti di misura erano distribuiti più uniformemente sul territorio. Tale differenza è dovuta chiaramente ai diversi obiettivi delle due campagne.

In questa nuova campagna, inoltre, non sono stati campionati solo locali siti al piano terra, come nella precedente, ma anche unità immobiliari situate ai piani superiori (Fig. 1-b).

La Tab. 3 mostra le principali statistiche descrittive relative alla concentrazione media annuale (ottenuta come media pesata dei due valori semestrali sull'effettivo periodo di esposizione) rilevata sui dati della campagna 2009/2010 e di quella precedente. Occorre precisare che un confronto diretto dei valori delle statistiche delle due campagne risulta inappropriato. Una prima differenza è dovuta al fatto che i punti della prima campagna erano situati esclusivamente al pian terreno diversamente da quella nuova in cui sono stati misurati anche i piani superiori. A questo proposito, nella Tab. 3 sono state riportate anche le statistiche relative ai soli piani terra misurati nella nuova campagna. Ciò che si osserva è un valore di concentrazione media annuale molto più elevato rispetto alla prima indagine (249 vs 124 Bq/m<sup>3</sup>), dovuto alla seconda differenza presente tra le due campagne: nell'indagine 2009/2010 il campionamento prevedeva, infatti, una percentuale maggiore di punti di misura in unità immobiliari situate in comuni appartenenti alla cosiddetta "fascia alta". Era, quindi, atteso un valore di concentrazione medio più elevato.



**Figura 1.** (a) Localizzazione dei punti di misura e suddivisione per tipologia di campionamento; (b) Ripartizione dei punti di misura per piano

Provincia	Tipologia a) (a1 e a2)	Tipologia b)	Tipologia c)	Tot
Bergamo	196	79	-	275
Brescia	113	44	25	182
Lodi	66	-	-	66
Milano	33	12	23	68
Mantova	113	-	-	113
Sondrio	154	35	26	215
Varese	46	1	-	47
<b>Tot</b>	<b>721</b>	<b>171</b>	<b>74</b>	<b>966</b>

**Tabella 1.** Numero di punti di misura previsti ed effettivi per provincia e tipologia di campionamento

	Campagna 2009 – 2010		Campagna 2003 - 2004	
	N° punti di misura	% punti di misura	N° punti di misura	% punti di misura
Fascia alta	709	73%	1126	32%
Fascia medio-alta	13	1%	768	22%
Fascia media	133	14%	423	12%
Fascia medio-bassa	0	0%	553	16%
Fascia bassa	111	12%	642	18%
Totale	966		3512	

**Tabella 2.** Ripartizione dei punti di misura per fascia e campagna di monitoraggio

Statistica	Campagna 2009-2010	Campagna 2009-2010 (solo punti al piano terra)	Campagna 2003-2004
Media	184	249	124
Dev. Standard	262	300	141
Mediana	90	132	78
Max	1936	1754	1796
% sopra i 100 Bq/m <sup>3</sup>	46%	62%	39%
% sopra i 200 Bq/m <sup>3</sup>	23%	35%	15%
% sopra i 400 Bq/m <sup>3</sup>	10%	17%	4%

**Tabella 3.** Statistiche descrittive della concentrazione media annuale di radon indoor suddivise per campagna di misura

## VALUTAZIONE DELLA RAPPRESENTATIVITA' DELLA CAMPAGNA 2003/2004

Uno degli obiettivi (obiettivo 1) della campagna 2009/2010 consiste nella valutazione della rappresentatività dei dati della precedente campagna di monitoraggio nella quale le unità immobiliari misurate non erano state estratte casualmente, ma perlopiù reperite su base volontaria ed erano localizzate esclusivamente al piano terra.

Per poter perseguire questo obiettivo è stato predisposto un piano di campionamento specifico (punti di tipologia a1 e a2) che prevedeva in primo luogo la scelta di alcuni comuni (per i quali si disponeva di un numero sufficiente di punti di misura della campagna 2003/2004) e l'estrazione casuale di famiglie dall'anagrafe (senza vincolo sul piano). In particolare, sono stati scelti 22 comuni, di cui il 55% classificato come di fascia alta, il 25% di fascia media e il 20% di fascia bassa. La rappresentatività dei dati della campagna 2003/2004 è stata valutata considerando due aspetti: in prima analisi i valori di concentrazione di radon indoor e successivamente la composizione del campione sulla base di alcune variabili legate alle caratteristiche di edificio o del territorio.

### Confronto delle due campagne regionali in termini di concentrazione di radon indoor

In prima battuta si è proceduto al confronto, per i comuni considerati, dei valori di concentrazione di radon indoor rilevati nelle due indagini. Per effettuare tale confronto è stato necessario scremare preliminarmente i dati, selezionando dal database 2003/2004 i soli punti dislocati nei 22 comuni di tipologia a1 e a2 (215 punti di misura) e dal database 2009/2010 i soli piani terra misurati in quei comuni (anche in questo caso 215 punti). Per due dei comuni analizzati non si dispone di un numero di misure sufficiente per ottenere stime affidabili e, per tal motivo, si è deciso di escluderli dal confronto considerando solo i 20 comuni restanti. Il confronto dei valori di concentrazione di radon indoor misurati nelle due campagne, singolarmente per ciascuno dei 20 comuni in analisi, è stato

effettuato tramite il test di Wilcoxon, il quale permette di verificare l'ipotesi di uguaglianza delle mediane di concentrazione comunale ottenute a partire dai dati delle due indagini. I risultati (Tab. 4) mostrano che per tutti i comuni viene accettata questa ipotesi ( $p\text{-value} > 0,05$  per tutti i comuni con l'eccezione di Mesenzana dove comunque il  $p\text{-value}$  è maggiore di 0,01).

Comune	Prov.	Fascia	N° di punti (2003 - 2004)	N° di punti (2009 - 2010)	Mediana (2003 -2004)	Mediana (2009 - 2010)	P-value
Cavacurta	LO	Bassa	12	9	54	67	0,2773
Corno Giovine	LO	Bassa	12	7	30	48	0,2614
Guardamiglio	LO	Bassa	14	6	39	42	0,659
Sermide	MN	Bassa	9	9	38	44	0,6665
Grone	BG	Media	5	14	87	120	0,7539
Castel Mella	BS	Media	9	13	77	72	0,7938
Marcaria	MN	Media	10	9	64	91	0,8421
Monzambano	MN	Media	6	8	108	58	0,1419
Virgilio	MN	Media	10	5	81	102	1
Castione	BG	Alta	10	23	290	460	0,4513
Clusone	BG	Alta	7	19	206	155	0,7777
Caino	BS	Alta	10	11	121	231	0,2816
Odolo	BS	Alta	10	11	176	178	0,8633
Castano Primo	MI	Alta	7	12	178	143	0,6504
Campodolcino	SO	Alta	10	11	205	221	0,9725
Piuro	SO	Alta	9	15	167	142	0,9534
Samolaco	SO	Alta	19	15	190	132	0,5836
Mesenzana	VA	Alta	7	4	121	64	0,0424
Saltrio	VA	Alta	8	5	189	48	0,3543
Viggiù	VA	Alta	10	6	109	137	0,7925

**Tabella 4.** Risultati del test di Wilcoxon per i comuni di tipologia a1-a2

## La rappresentatività in termini di composizione del campione

Un secondo obiettivo è stato quello di valutare la rappresentatività della campagna 2003/2004 in termini di composizione del campione sulla base di alcune variabili legate, per esempio, alle caratteristiche di edificio o del territorio. L'idea alla base di questa analisi si origina dalla considerazione che gli edifici della campagna 2009/2010, essendo stati estratti casualmente, possono considerarsi un campione rappresentativo della popolazione di edifici nei comuni considerati, al contrario dei locali indagati nella campagna 2003/2004 che appartenevano ad unità immobiliari perlopiù individuate su base volontaria. Se il campione della campagna 2003/2004 presenta una composizione simile a quello della campagna 2009/2010 rispetto ad una particolare variabile, è plausibile ritenere anche la prima indagine rappresentativa rispetto a quell'attributo.

Nello specifico, le variabili che sono state prese in considerazione sono:

- anno di costruzione/ultima ristrutturazione effettuata;
- caratteristiche dell'edificio, rispetto al piano campagna;
- uso del suolo nelle aree urbanizzate (tali informazioni sono state ottenute effettuando una sovrapposizione in GIS tra i punti di misura e la Carta regionale dell'uso del suolo<sup>1</sup>);

Tramite test statistici atti a verificare l'uguaglianza fra la proporzione di unità aventi una particolare caratteristica in due gruppi (Agresti, 2002) si è ottenuto, per ogni variabile considerata, che le due campagne presentano una composizione non significativamente diversa. Si può, quindi, ritenere che

<sup>1</sup> La carta è reperibile sul sito "Geoportale della Lombardia": <http://www.cartografia.regione.lombardia.it/geoportale>.

la campagna 2003/2004 sia rappresentativa della zona indagata rispetto alle variabili considerate. Valendo ciò su tutte le porzioni di territorio in cui la verifica poteva essere fatta (comuni campionati di tipo a1 e a2), possiamo ipotizzare che questo possa mantenersi valido anche per l'intera regione, sebbene sia stato possibile verificare direttamente quest'ipotesi solo su un'area ristretta del territorio lombardo.

## LA CONCENTRAZIONE DI RADON INDOOR AL VARIARE DEL PIANO

Come ricordato in precedenza, un'ulteriore obiettivo della campagna 2009/2010 consiste nel valutare come varia la concentrazione di radon indoor al variare del piano del locale di misura. Per questo obiettivo sono stati campionati i punti di tipologia a1 e a2. A questi soli punti si è fatto riferimento nelle analisi riportate in questa sezione.

La principale sorgente di radon, come è noto, è rappresentata dal suolo e nello specifico dalla natura delle rocce di cui è costituito e dal loro contenuto di radioattività naturale. In linea di principio è, quindi, atteso avere maggiori concentrazioni in locali più vicini al terreno, come quelli situati al piano terra o, ancor di più, interrati o seminterrati. Questo aspetto emerge già a livello descrittivo osservando le concentrazioni medie per piano riportate in Tab. 5.

Numerosi studi hanno evidenziato l'impatto del piano in cui il locale è situato sulle misurazioni di radon all'interno dello stesso (si veda per esempio Antignani *et al.*, 2009; De Francesco *et al.*, 2010; Franco-Marina *et al.*, 2003; Gunby *et al.*, 1993; Price *et al.*, 1996; Smith e Field, 2007; Sundal *et al.*, 2004; Zhu *et al.*, 2001).

In prima analisi, per valutare se sussistono differenze significative di concentrazione tra piani diversi si è proceduto effettuando il test di Kruskal e Wallis che verifica l'ipotesi che congiuntamente le mediane per piano siano tutte uguali tra loro. Il risultato del test porta al rifiuto di tale ipotesi indicando che esiste almeno una coppia di piani per cui le mediane risultano significativamente diverse. Per capire quale/i coppie di piani risultano statisticamente diverse sono stati effettuati dei confronti a coppie tramite il test di Wilcoxon<sup>2</sup>. I risultati dei test mostrano una differenza statisticamente significativa tra i piani terra confrontati con i piani rialzati, i primi e i secondi e tra i piani seminterrati e rialzati. Non sussistono differenze significative per i confronti tra primi, secondi e terzi piani, indicando quindi che la concentrazione in tali piani è pressoché simile<sup>3</sup>.

Piano di misura	N° punti di misura	Media	Dev. Standard	Mediana	% locali sopra i 200 Bq/m <sup>3</sup>	% locali sopra i 400 Bq/m <sup>3</sup>
I	4	201	128	161	25%	0%
S	14	390	509	197	50%	36%
T	215	233	307	114	30%	16%
R	30	76	90	43	10%	3%
1	287	123	206	68	12%	3%
2	129	134	226	70	14%	7%
3	36	165	199	84	22%	11%
4	5	115	143	53	20%	0%
5	1	195	-	195	0%	0%
Tot	721					

**Tabella 5.** Statistiche descrittive della concentrazione di radon indoor stratificate per piano (punti di tipologia a1-a2)

Un'analisi più precisa per valutare l'effetto del piano sulla concentrazione di radon indoor, può essere condotta ricorrendo ad opportuni modelli di regressione che consentono di valutare tale effetto al netto di quello di altri fattori potenzialmente rilevanti. Tale approccio è stato seguito,

<sup>2</sup> Poiché sono stati effettuati confronti multipli i p-value associati ai test sono stati opportunamente aggiustati tramite la correzione di Bonferroni.

<sup>3</sup> Occorre comunque osservare che i risultati dei confronti a coppie che coinvolgono piani interrati, seminterrati, quarti o quinti potrebbero essere influenzati anche della scarsa numerosità di ciascuno di questi gruppi.

utilizzando modelli diversi, da numerosi autori (Smith e Field, 2007; Gunby *et al.*, 1993; Price *et al.*, 1996), sebbene altri metodi siano stati proposti in letteratura. Per esempio, al fine di normalizzare misure rilevate a piani diversi allo scopo di proporre mappe di concentrazione per il radon indoor, Chiaberto *et al.* (2006) hanno suggerito di stimare il rapporto tra la concentrazione ad un determinato piano e quella rilevata in un edificio al piano terra tramite il rapporto fra le medie geometriche delle misure rilevate ai quei piani. Tale procedura si basa sull'assunto che esista una relazione di proporzionalità diretta fra il piano e la concentrazione indoor del radon e che quest'ultima segua, al piano terra e quindi anche a tutti gli altri piani, una legge log-normale di parametri  $\mu$  e  $\sigma$ . Quest'approccio non è però applicabile ai dati della campagna lombarda in quanto la relazione di linearità fra piano del locale di misura e concentrazione rilevata non risulta un'ipotesi coerente, come evidente anche dalla Tab. 5. Inoltre, la metodologia proposta, essendo di natura univariata, non consente di controllare l'effetto di altri fattori influenti che potrebbero alterare tale relazione. Infine, sebbene il logaritmo della media geometrica sia un buon stimatore del parametro  $\mu$  (esso risulta essere infatti uno stimatore corretto e, in quanto stimatore di massima verosimiglianza, dotato di buone proprietà asintotiche), la media geometrica può risultare molto distorta, soprattutto in campioni non particolarmente numerosi, come stimatore del valore atteso della concentrazione, anche nel caso in cui l'assunzione di linearità sia valida. In questo lavoro si è quindi preferito seguire l'approccio di regressione ricordato in precedenza.

L'approccio usuale per trattare problemi di questo tipo è applicare un modello di analisi della varianza (ANOVA) ai valori delle concentrazioni trasformati tramite logaritmo. Indicata con  $Y$  la concentrazione di radon indoor, con  $P$  la variabile categoriale rappresentante il piano del locale di misura strutturata genericamente su  $k$  livelli e con  $\mathbf{x}$  un vettore di variabili potenzialmente influenti sulle concentrazioni, il modello in questione (modello 1) può essere formalizzato come:

$$(1) \quad E(\log Y | P = j, \mathbf{x}) = \alpha + \lambda_j + \mathbf{x}'\boldsymbol{\beta}$$

dove l'operatore  $E(\log Y | P = j, \mathbf{x})$  rappresenta il valore atteso di  $\log Y$  condizionato al  $j$ -esimo piano del locale e ai valori di  $\mathbf{x}$ , mentre  $\alpha$ ,  $\boldsymbol{\beta}$ ,  $\lambda_j$ ,  $j=1, \dots, k$ , sono parametri incogniti che misurano l'effetto delle corrispondenti caratteristiche dell'edificio. Data l'assunzione di log-gaussianità, spesso ragionevole per le concentrazioni di radon indoor (Nazaroff e Nero, 1988), la trasformazione logaritmica appare adeguata anche per garantire gli assunti di normalità su cui si basano molte procedure inferenziali legate al modello lineare sopra specificato. Il problema che può presentare il modello nella (1) è di fornire dei parametri d'effetto validi sulla scala logaritmica. Spesso, invece, occorre disporre di stime valide sulla scala originale (ad esempio per avere dei fattori di conversione tra le misurazioni ricavate a piani diversi o stimare le concentrazioni per particolari profili di locale). L'usuale prassi, in questo caso, è quella di esponenziare le stime ottenute per ricondurle a tale scala. Evidentemente però, essendo  $E(e^Y) \neq e^{E(Y)}$ , tale trasformazione induce delle distorsioni sistematiche nell'inferenza. In particolare, per la disuguaglianza di Jensen, si tenderanno ad avere sottostime sistematiche. Un modo per risolvere questo tipo di problema è quello di estendere il modello lineare (1) ad una classe nota come modelli lineari generalizzati (GLM, Myers *et al.*, 2010).

Una prima possibilità è, mantenendo l'assunto di normalità, di specificare un modello (modello 2) del tipo

$$(2) \quad \log(E(Y | P = j, \mathbf{x})) = \alpha + \lambda_j + \mathbf{x}'\boldsymbol{\beta}.$$

In questo caso la trasformazione logaritmica (nota nella letteratura sui GLM come *link function*) lega il valore atteso di  $Y$  alla componente lineare. Questo consente di risolvere il problema della distorsione menzionato in precedenza. Tuttavia ipotizzare delle concentrazioni normalmente distribuite sembra poco consistente con l'osservazione empirica tipica delle concentrazioni del radon indoor. Un valido compromesso risulta quello di mantenere la specificazione (2), ma introdurre una relazione di proporzionalità tra la media e la varianza di  $Y$  che è tipica della legge log-normale. In questo caso la stima del modello viene ottenuta tramite un metodo noto come quasi-verosimiglianza (modello 3). Per i dettagli si rimanda a Myers *et al.*, 2010 pag. 211 e seguenti.

I tre modelli sopra ricordati sono stati tutti applicati al dataset descritto nella sezione precedente e l'approccio di quasi-verosimiglianza è risultato il migliore sia in termini di analisi dei residui (non riportata per esteso nel presente lavoro) che in base ad un'analisi di cross validazione in cui, seguendo una procedura leave-one-out, si è valutata la capacità predittiva di ognuno dei tre modelli considerati in termini di errore quadratico medio (EQM), calcolato come  $\sum_{i=1}^n (y_i - \hat{y}_{i(-i)})^2 / n$  dove  $n$  rappresenta la numerosità del campione,  $y_i$  il valore della concentrazione rilevato nell' $i$ -esimo punto di misura campionario e  $\hat{y}_{i(-i)}$  il valore della concentrazione stimato per l' $i$ -esimo punto di misura rimuovendo  $y_i$  dal dataset e utilizzando i rimanenti  $n-1$  per la stima.

Modello	$R^2$	EQM
Modello 1	0,22	54590
Modello 2	0,24	50579
Modello 3	0,23	49811

**Tabella 6.**  $R^2$  e errore quadratico medio (EQM) ottenuto tramite cross validazione leave-one-out per i 3 modelli

Poiché quest'ultimo modello presenta il più piccolo valore di EQM, come risulta evidente dalla Tab. 6 in cui vengono mostrati anche i valori di  $R^2$ , si è deciso di riportare in questo lavoro solo i risultati relativi ad esso. Più precisamente quali variabili esplicative, oltre al piano (accorpato in 5 categorie: interrato/seminterrato, terra, rialzato/1°, 2°, 3° o più), si sono considerate: la fascia del comune (bassa, media, alta), la tipologia di edificio (isolato o contiguo su almeno un lato con altri edifici), il contatto con il terreno (a diretto contatto o isolato tramite vespaio o sotterraneo), le ore giornaliere di apertura delle finestre del locale di misura (minore o maggiore di 6 ore), l'anno di costruzione dell'edificio (prima del 1970 o dopo) e il materiale delle pareti (pietra o altro materiale). Dalle stime dei parametri per il modello considerato, riportate in Tab. 7, si osserva come locali situati in edifici isolati, a contatto con il terreno, costruiti prima del 1970 con meno di 6 ore giornaliere di apertura della finestre tendano a presentare dei valori di concentrazione mediamente più elevati. Analogamente le concentrazioni rilevate in locali situati ai piani bassi, in particolare dagli interrati al primo piano, tendono ad essere mediamente più elevate di quelle presenti ai piani più alti. A partire dal secondo piano, invece, la concentrazione di radon nel locale risulta sostanzialmente non influenzata dal piano in cui il locale è situato. Nel modello è stata inclusa anche una variabile categoriale rappresentante la fascia territoriale a cui il comune appartiene. Si è introdotta questa variabile informativa del contesto ambientale per avere stime dei parametri relativi all'effetto del piano, depurate, il più possibile, da tale componente. Le stime ottenute per i parametri associati a ciascuno dei tre livelli confermano, come era atteso, la classificazione utilizzata.

Il modello stimato consente, infine, di ottenere dei "coefficienti di piano", ossia dei coefficienti moltiplicativi che possono essere applicati alla concentrazione del piano terra per ottenere quella di un determinato piano. Tali coefficienti sono riportati in Tab. 7. L'intercetta del modello fornisce la stima della concentrazione di radon indoor al piano terra per un locale di riferimento (cioè un locale con pareti non in pietra e con apertura frequente delle finestre, posto in un edificio sito in un comune di fascia bassa, non isolato, non a diretto contatto con il terreno e costruito dopo il 1970). Esponenziando l'intercetta si ottiene il valore di concentrazione previsto al piano terra per un locale di riferimento, ovvero  $\exp(3,84) = 46,34 \text{ Bq/m}^3$ . Per ottenere una stima della concentrazione in un locale situato in un edificio con le medesime caratteristiche, ma ad un piano diverso, per esempio interrato o seminterrato, è necessario moltiplicare la concentrazione stimata al piano terra per il coefficiente del piano considerato ottenuto esponenziando il coefficiente di regressione corrispondente. Per un locale al piano interrato/seminterrato questo coefficiente è, ad esempio, pari a  $1,65 = \exp(0,5)$ , da cui discende che la concentrazione prevista a quel piano è pari a  $46,34 \times 1,65 = 76,5 \text{ Bq/m}^3$ . La concentrazione ad un piano interrato/seminterrato, quindi, aumenta mediamente del 65% rispetto a quella di un locale al piano terra. In un locale al 2° piano diminuisce



(essendo il coefficiente inferiore a 1) invece del 52%. Tale diminuzione può variare dal 41% al 66%, come visibile dall'intervallo di confidenza associato.

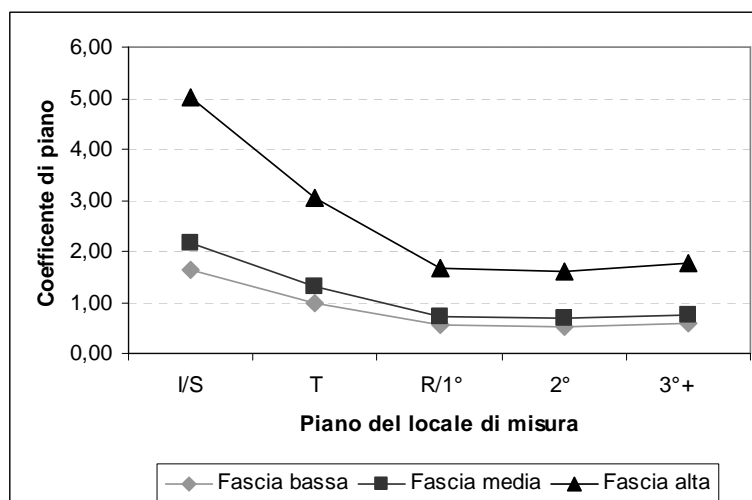
In Fig. 2 è stata rappresentata mediante una spezzata la relazione piano-coefficiente stimata dal modello, per ognuna delle 3 fasce.

Si osserva infine che, in modo analogo a quanto discusso per il piano, è possibile derivare dei coefficienti moltiplicativi anche per le altre variabili incluse nel modello anch'essi riportati in Tab. 7. Questo consente di ottenere le concentrazioni previste per locali situati a diversi piani e con diverse caratteristiche. Se ad esempio, si vuole ottenere la concentrazione stimata per un locale al 2° piano in un comune di fascia alta, posto in un edificio isolato e a contatto con il terreno, costruito prima del 1970, occorre moltiplicare il valore di concentrazione previsto al piano terra per un locale di riferimento, 46,34 Bq/m<sup>3</sup>, per i coefficienti relativi a queste ultime variabili, e quindi 0,52 (coefficiente del 2° piano), 3,06 (coefficiente relativo alla fascia alta), 1,22 (coefficiente relativo ad un edificio isolato), 1,27 (coefficiente relativo ad un edificio a contatto con il terreno), 1,49 (coefficiente relativo ad un edificio costruito prima del 1970). Questo equivale ad applicare un fattore moltiplicativo complessivo pari a 3,67.

	Stima	Standard error	P-value	Coefficienti moltiplicativi	IC del coefficiente moltiplicativo	
					Inf	Sup
Intercetta	3,84	0,14	<0,001	1	-	-
I/S	0,50	0,27	0,068	1,65	0,96	2,81
R/1°	-0,60	0,09	<0,001	0,55	0,46	0,66
2°	-0,65	0,12	<0,001	0,52	0,41	0,66
3°+	-0,55	0,18	0,002	0,58	0,41	0,82
Fascia media	0,28	0,15	0,060	1,32	0,99	1,76
Fascia alta	1,12	0,12	<0,001	3,06	2,40	3,89
Locale isolato	0,20	0,09	0,021	1,22	1,03	1,44
Locale a contatto con il terreno	0,24	0,08	0,004	1,27	1,08	1,49
Ore di apertura della finestre < 6 h	0,17	0,08	0,041	1,19	1,01	1,40
Edificio costruito prima del 1970	0,40	0,09	<0,001	1,49	1,25	1,76
Pareti in pietra	0,46	0,11	<0,001	1,59	1,28	1,98

\* Categorie di riferimento per le variabili nel modello: Piano: Terra; Fascia: Bassa; Tipologia di edificio: Locale non isolato; Piano Campagna: Locale non a contatto con il terreno; Ore di apertura giornaliera delle finestre: > 6 h; Anno di costruzione: dopo il 1970; Pareti: altro materiale diverso dalla pietra.

**Tabella 7.** Stime del modello di regressione e coefficienti di piano



**Figura 2.** Relazione piano-coefficiente stimata dal modello per fascia

## CONCLUSIONI

Nel presente lavoro è stata fornita una descrizione della seconda campagna di monitoraggio della concentrazione di radon indoor realizzata nel 2009/2010 in Lombardia, presentandone gli obiettivi e il piano di campionamento effettuato per reperire le unità immobiliari da misurare.

Tra i diversi obiettivi di indagine, in questo lavoro sono stati considerati i risultati relativi a due di essi. In primo luogo si è cercato di valutare la rappresentatività della precedente campagna di monitoraggio (2003/2004) nella quale i punti di misura, situati esclusivamente al pian terreno, non erano stati estratti casualmente, differentemente dalla nuova campagna in cui il campione è stato ottenuto estraendo casualmente le famiglie dagli elenchi dell'anagrafe. Le analisi effettuate hanno mostrato che la precedente indagine può considerarsi rappresentativa, limitatamente all'insieme dei locali posti a pian terreno, sia in termini di valori di concentrazione che in termini di composizione del campione rispetto a specifiche variabili. Secondariamente, è stata effettuata una valutazione di come varia la concentrazione al variare del piano del locale di misura. Come era atteso, si sono osservate concentrazioni maggiori ai piani più vicini al suolo, essendo proprio quest'ultimo la principale sorgente di radon. Poiché questo fenomeno è influenzato da altri numerosi fattori legati sia a caratteristiche ambientali che di edificio, è stato implementato un opportuno modello di regressione in modo da valutare l'effetto del piano, al netto di altri fattori influenti.

I risultati relativi agli altri tre obiettivi, non riportati in questo lavoro, sono descritti in un dossier dettagliato consegnato alla Regione Lombardia a luglio 2011.

## BIBLIOGRAFIA

- Agresti A. (2002), *Categorical Data Analysis*, John Wiley & Sons, New York
- Antignani S., Bochicchio F., Ampollini M., Venoso G., Bruni B., Innamorati S., Malaguti L., Stefano A. (2009), Radon concentration variations between and within buildings of a research institute, *Radiation Measurements*, 44, 1040–1044
- Borgoni R., Cremonesi A., Somà G., de Bartolo D., Alberici A. (2007), *Radon in Lombardia: dai valori di concentrazione indoor misurati, all'individuazione dei Comuni con elevata probabilità di alte concentrazioni. Un approccio geostatistico*, AIRP- Atti del XXXIV Congresso Nazionale di Radioprotezione (Vasto, 1-3 ottobre 2007)
- Chiaberto E., Magnoni M., Righino F., Serena E., Procopio S. (2006), *Il radon in Piemonte. Dalla distribuzione delle concentrazioni alla definizione delle aree a rischio: criteri a confronto*. AIRP - Atti del XXXIII Congresso Nazionale di Radioprotezione (Torino, 20 - 23 settembre 2006)
- De Francesco S., Pascale Tommasone F., Cuoco E., Tedesco D. (2010), Indoor radon seasonal variability at different floors of buildings, *Radiation Measurements*, 45, 928-934
- Franco-Marina F., Villalba-Caloca J., Segovia N., Tavera L. (2003), Spatial indoor radon distribution in Mexico City, *The Science of the Total Environment*, 317, 91-103
- Gunby J.A., Darby S.C., Miles J.C., Green B.M., Cox D.R. (1993), Factors affecting indoor radon concentrations in the United Kingdom, *Health Physics*, 64, 2-12
- Ministero della Salute, (2002), Piano Nazionale Radon.
- Myers R.H., Montgomery D.C., Vining G.G., Robinson T.J. (2010), *Generalized Linear Models: with applications in engineering and the science*, 2<sup>nd</sup> ed, John Wiley and Sons
- Nazaroff W.W., Nero A.V. (1988), *Radon and its decay products in indoor air*, John Wiley and Sons, New York
- Price P.N., Nero A.V., Gelman A. (1996), Bayesian prediction of mean indoor radon concentrations for Minnesota counties, *Health Physics*, 71, 922-936
- Smith B.J., Field R.W. (2007), Effect of housing factor and sufficient uranium on the spatial prediction of residential radon in Iowa, *Environmetrics*, 18, 481–497
- Sundal A.V., Henriksen H., Soldal O., Strand T. (2004), The influence of geological factors on indoor radon concentrations in Norway, *The Science of the Total Environment*, 328, 41-53
- Zhu H.C., Charlet J.M., Poffijn A. (2001), Radon risk mapping in southern Belgium: an application of geostatistical and GIS techniques, *The Science of the Total Environment*, 272, 203-210