

Analisi preliminare per la rappresentatività spaziale delle stazioni di monitoraggio di qualità dell'aria delle polveri nella Val d'Agri – Basilicata

Anna Maria Crisci, Lucia Mangiamele, Raffaella Esposito, Michele Lovallo
ARPAB, Potenza - 85100

Nell’ambito delle attività di predisposizione del Progetto di Adeguamento della rete di misura e del Programma di valutazione, di cui all’art.5 del D.Lgs. 155/2010 e ss.mm.ii., sono stati avviati studi statistici dei dati di PM10 e di PM2.5, rilevati dalla rete regionale di monitoraggio della qualità dell’aria della Basilicata di ARPAB, per esaminarne la rappresentatività spaziale. Inoltre, in funzione di una ottimizzazione della rete si intende individuare le stazioni che esprimono misure ridondanti. La rete regionale della qualità dell’aria dell’ARPAB (Fig.1) è costituita da 15 centraline di differente classificazione e tipologia, per sensoristica installata e caratteristiche dell’area di installazione. Ai fini della presente analisi, è stata presa in considerazione la sottorete di monitoraggio della zona della Val d’Agri (VDA), in prossimità del COVA (Centro Olio Val d’Agri), costituita dalle quattro stazioni dotate di analizzatori di misura automatica di PM10 e PM2.5 ad assorbimento β . La base di dati utilizzata nello studio è costituita dalla serie storica delle concentrazioni giornaliere di PM10 e PM2.5 sull’arco temporale che va dal 1 marzo 2013 al 1 marzo 2018.



Mappa della rete di monitoraggio della Basilicata



Dettaglio della rete di monitoraggio in Val d’Agri

Nome Stazione	Coordinate Piane (m)		Quota Ortometrica (m s.l.m)
	(sistema ETRS89)		
	E	N	Q
Viggiano 1	576507	4465314	820
Grumento 3	575772	4460148	735
Masseria De Blasiis	573687	4464106	603
Costa Molina Sud 1	581126	4463226	690

Coppia di Stazioni	Distanza (m)
Viggiano 1- Costa Molina Sud 1	5071
Viggiano 1 - Masseria De Blasiis	3076
Viggiano 1 - Grumento 3	5219
Costa Molina Sud 1 - Masseria De Blasiis	7491
Costa Molina Sud 1 – Grumento 3	6176
Masseria De Blasiis – Gruemento 3	4476

Metodo

L'elaborazione statistica è stata effettuata applicando differenti metodi statistici allo scopo di individuare quelli più immediati ed efficaci nel fornire risultati utili alla valutazione della rappresentatività spaziale delle stazioni rispetto ai parametri "concentrazione di PM10", "concentrazione di PM2.5" e "rapporto della concentrazione di PM10 su PM2.5". I metodi statistici utilizzati sono stati in particolare:

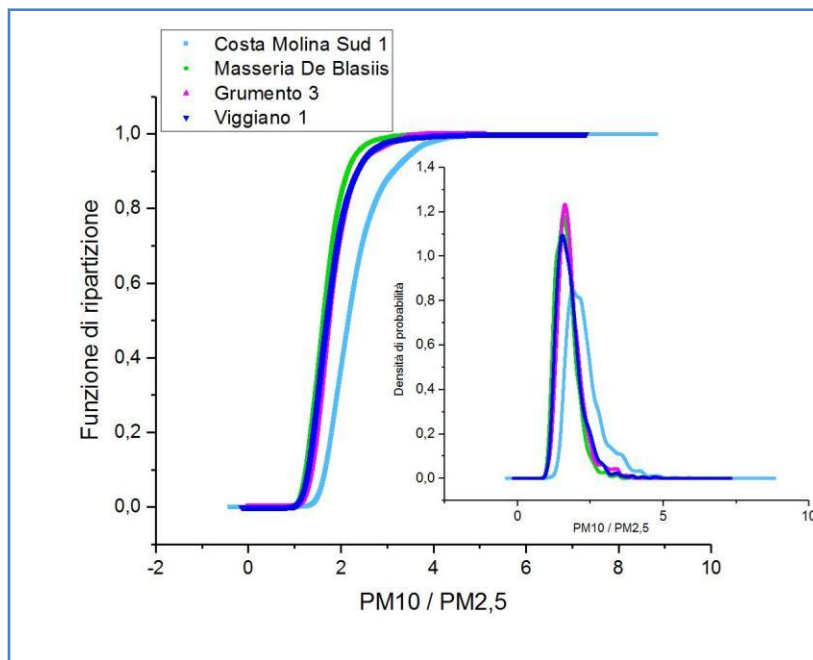
- ❖ la determinazione della distribuzione di probabilità del PM10, del PM2.5 e del loro rapporto applicando il metodo non parametrico del Kernel;
- ❖ i box-plot e i bar-Plot, per la determinazione dei principali parametri standard della statistica descrittiva;
- ❖ l'analisi della correlazione tra le serie di dati del PM10 e PM2.5 per ogni coppia di stazioni e della correlazione tra le coppie di concentrazioni di particolato per ogni stazione;
- ❖ valutazione del grado di accuratezza e di precisione delle precedenti correlazioni tramite il calcolo del coefficiente di Pearson e del coefficiente di concordanza;
- ❖ l'analisi delle componenti principali PCA.

Analisi

Distribuzione di densità di probabilità

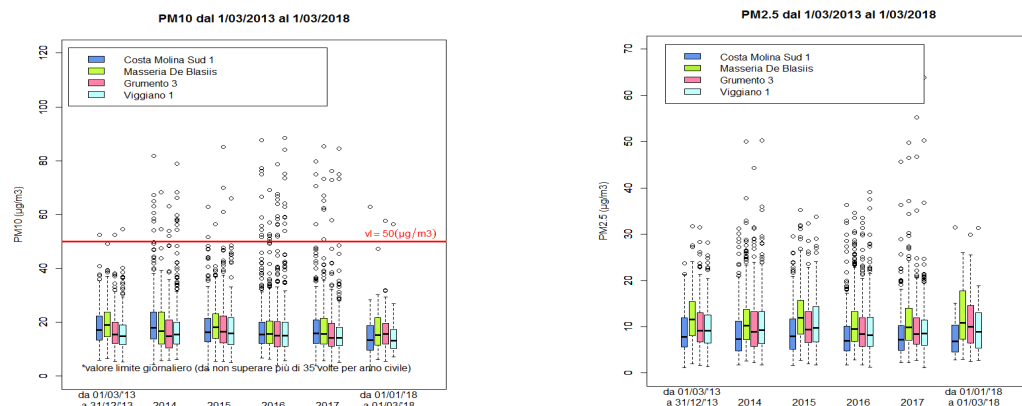
Le distribuzioni di densità di probabilità e le relative funzioni di ripartizione del PM10 possono ritenersi sostanzialmente sovrapposte per tutte le stazioni; per il PM2.5, invece, le distribuzioni di Viggiano 1 e Grumento 3 risultano pressoché identiche, mentre si notano delle differenze per le stazioni di Costa Molina Sud 1 e Masseria De Blasiis. Le concentrazioni sia di PM10 sia di PM2.5 con maggiore occorrenza sono quelle che ricadono rispettivamente intorno ai $20 \mu\text{g}/\text{m}^3$ e ai $10 \mu\text{g}/\text{m}^3$.

Per meglio evidenziare le differenze tra PM10 e PM2.5 tra le varie stazioni, si sono determinate le distribuzioni delle densità di probabilità del rapporto PM10/PM2.5 dove l'andamento di quella relativa alla stazione di Costa Molina Sud 1, pur mantenendo la stessa forma, risulta traslato verso valori più elevati e con un massimo sensibilmente più basso. A titolo di esempio, si riportano i grafici del rapporto PM10/PM2.5.



Box-plot

Dall'analisi dei box-plot (cfr. tabelle in calce) emerge un valore medio di concentrazione, sull'intero periodo 2013-2018, pressoché costante tra le diverse stazioni, pari circa a 18 $\mu\text{g}/\text{m}^3$ per il PM10 ed a 10 $\mu\text{g}/\text{m}^3$ per il PM2.5. E' peraltro evidente che le concentrazioni medio-annue di PM2.5 più alte sono registrate a Masseria De Blasiis, mentre quelle più basse a Costa Molina, a differenza del PM10 che si mantiene costante tra i diversi siti di misura. Con riferimento ai valori limite del PM10, non si misurano superamenti sia del valore limite giornaliero (n. 35 superamenti di 50 $\mu\text{g}/\text{m}^3$) che del limite annuo (40 $\mu\text{g}/\text{m}^3$). Relativamente ai dati giornalieri di PM10, nei casi in cui si è verificata la contemporaneità degli incrementi di concentrazione su diverse stazioni della rete di monitoraggio, è presumibile che la causa sia legata al trasporto di polveri desertiche e successiva intrusione a livello locale.



L'estremo inferiore, quello superiore e la linea orizzontale interna dei box presenti nei grafici, rappresentano rispettivamente il primo percentile (Q1), il terzo percentile (Q3) e la mediana (Q2) della serie considerata; in aggiunta sono riportati, esternamente ai box, i "baffi" che consistono in linee verticali tratteggiate delimitate dai "valori adiacenti inferiore e superiore" (VAI e VAS):

VAI è il più piccolo valore che risulta $\geq Q1 - 1.5 * r$

VAS è il più grande valore che risulta $\leq Q3 + 1.5 * r$

$r = Q3 - Q1$ è la distanza interquantilica

Esternamente ai baffi vengono individuati i punti outliers.

PM10 ($\mu\text{g}/\text{m}^3$)		Viggiano 1	Costa Molina Sud1	Grumento 3	Masseria De Blasiis
2013	Media	15,87	18,33	16,72	19,83
	Mediana	14,81	17,18	15,4	11,6
	Copertura	300	300	285	283
	Dati	(82%)	(82%)	(78%)	(78%)
2014	Media	17,96	20,26	17,25	18,93
	Mediana	15,5	17,9	14,9	10,30
	Copertura	341	357	205	203
	Dati	(93%)	(98%)	(56%)	(56%)
2015	Media	17,54	18,01	18,22	19,22
	Mediana	15,8	16,42	16,49	11,9
	Copertura	355	361	340	346
	Dati	(97%)	(99%)	(93%)	(95%)
2016	Media	17,93	18,09	17,4	17,72
	Mediana	15,1	15,55	15	9,5
	Copertura	358	351	356	343
	Dati	(98%)	(96%)	(97%)	(94%)
2017	Media	15,96	18,01	16,67	18,57
	Mediana	14,2	15,915	14,21	9,91
	Copertura	358	342	351	358
	Dati	(98%)	(94%)	(96%)	(98%)
2018	Media	14,63	15,45	17,04	16,68
	Mediana	13,3	13,38	15,7	10,8
	Copertura	59	53	59	58
	Dati	(16%)	(15%)	(16%)	(16%)

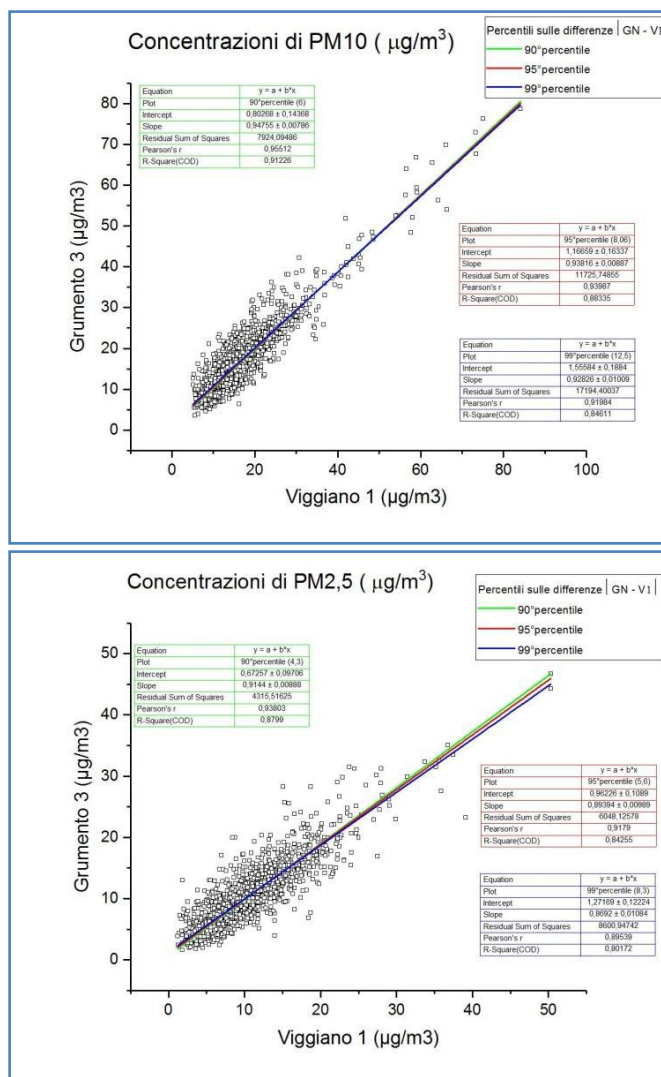
PM2,5 ($\mu\text{g}/\text{m}^3$)		Viggiano 1	Costa Molina Sud 1	Grumento 3	Masseria De Blasiis
2013	Media	9,88	8,86	10,13	12,06
	Mediana	9,2	7,81	9,2	11,6
	Copertura	300	259	281	293
	Dati	(82%)	(71%)	(77%)	(80%)
2014	Media	10,65	9,02	10,44	11,45
	Mediana	9,3	7,4	8,9	10,3
	Copertura	298	248	243	350
	Dati	(82%)	(68%)	(67%)	(96%)
2015	Media	10,85	8,88	10,21	12,73
	Mediana	9,8	7,9	9,4	11,9
	Copertura	357	358	353	354
	Dati	(98%)	(98%)	(97%)	(97%)
2016	Media	9,54	8,09	9,57	10,91
	Mediana	8,2	6,95	8,42	9,5
	Copertura	360	352	358	362
	Dati	(99%)	(96%)	(98%)	(99%)
2017	Media	9,51	8,31	9,68	11,2
	Mediana	8,6	7,2	8,4	9,91
	Copertura	355	351	355	354
	Dati	(97%)	(96%)	(97%)	(97%)
2018	Media	9,71	8	10,94	12,07
	Mediana	8,9	6,9	10	10,8
	Copertura	59	58	59	57
	Dati	(16%)	(16%)	(16%)	(16%)

Regressione lineare

L'analisi ha avuto seguito con le regressioni lineari sulle serie storiche di PM10 e PM2.5 per ogni coppia di stazione al fine di stabilirne il grado di correlazione. Dai risultati, si evince una buona correlazione tra le stazioni su entrambi i parametri, come si può vedere dai valori del coefficiente di Pearson ρ riportati nelle tabella successive. Per valutare l'influenza del contributo degli eventi outliers, si sono messe a punto due ulteriori tipologie di regressione:

1. il calcolo del fit è avvenuto dopo avere effettuato un taglio al 95° percentile sui dati delle serie storiche per coppie di stazioni. In tal caso, il taglio dei dati al 95° percentile implica l'eliminazione dei valori maggiori del dato di concentrazione corrispondente al 95% della serie ordinata di dati;
2. i fit sono stati eseguiti al 90°, al 95° e al 99° percentile sulle differenze tra le concentrazioni, nella stessa data, nella coppia di stazioni presa in esame. Agendo sulle differenze, il taglio non avviene necessariamente sugli outliers ma sullo scarto di concentrazione superiore rispettivamente al 90% - 95% - 99% della serie ordinata di dati e correlabile ad eventi di picco prettamente locali.

Ritenendo maggiormente significativi i risultati afferenti al secondo metodo all'interno del quale la correlazione è rafforzata dal contributo degli eventi outliers che risultano comuni a più siti, se ne riportano i grafici. In particolare la stazione di Viggiano 1 risulta fortemente correlata sia con Costa Molina Sud 1 sia con Grumento 3 considerando i fit al 90°, al 95° e al 99° percentile.



Coefficiente di Lin

In questa fase si è valutato il coefficiente di concordanza di Lin “CCC” (*Concordance correlation coefficient*) che dà indicazione della concordanza in termini di omogeneità spaziale.

Dalle tabelle si evince che:

1. le stazioni maggiormente concordanti sono Costa Molina Sud 1 e Grumento 3 con Viggiano 1;
2. la stazione di Masseria De Blasiis è meno concordante con Viggiano 1 e Costa Molina Sud 1.

PM10	ρ di Pearson	CCC
Costa Molina Sud 1 – Grumento 3	0,83	0,79
Costa Molina Sud 1 - Masseria De Blasiis	0,80	0,68
Costa Molina Sud 1 – Viggiano1	0,89	0,85
Grumento 3 - Viggiano1	0,89	0,89
Masseria De Blasiis - Grumento 3	0,85	0,81
Viggiano 1 – Masseria De Blasiis	0,86	0,83

PM2,5	ρ di Pearson	CCC
Costa Molina Sud 1 – Grumento 3	0,88	0,88
Costa Molina Sud 1- Masseria De Blasiis	0,79	0,79
Costa Molina Sud 1 – Viggiano1	0,89	0,88
Grumento 3 - Viggiano1	0,89	0,89
Masseria De Blasiis - Grumento 3	0,82	0,81
Viggiano 1 – Masseria De Blasiis	0,77	0,75

Analisi delle componenti principali

Come ulteriore elaborazione è stata effettuata l'*Analisi delle componenti principali* (PCA) alle serie di PM10 e di PM2.5 delle quattro stazioni, considerate come variabili originarie. Dai risultati è emerso che le componenti principali sono quattro per ogni serie: la prima componente principale della serie di PM10 spiega la varianza all'87.8% e la prima componente principale riferita alla serie di PM2.5 ne spiega l'88.8%.

Inoltre, tali componenti principali, combinazione lineare delle variabili originarie, presentano *pesi* (componenti degli autovettori corrispondenti agli autovalori scelti) con valori simili tra loro, come mostrato nella tabella seguente.

	Prima Componente Principale - PC1
PM2.5	PC1= 0.50 CMS1 + 0.49 MDB + 0.50 GRU3 + 0.51 VIG1
PM10	PC1= 0.51 CMS1 + 0.47 MDB + 0.51 GRU3 + 0.51 VIG1

(CMS1=Costa Molina Sud 1 =; MDB Masseria De Blasiis; GRU3= Grumento 3; VIG1=Viggiano 1)

Da questo quadro è possibile desumere che le quattro stazioni sono interscambiabili in termini di misure di PM10 e PM2.5.

Soglia di valutazione superiore ed inferiore

	Media su 24 ore PM10	Media annuale PM10	Media annuale PM2.5
Soglia di valutazione superiore	70% del valore limite (35 µg/m³ da non superare più di 35 volte per anno civile)	70% del valore limite (28 µg/m³)	70% del valore limite (17 µg/m³)
Soglia di valutazione inferiore	50% del valore limite (25 µg/m³ da non superare più di 35 volte per anno civile)	50% del valore limite (20 µg/m³)	50% del valore limite (12 µg/m³)

In ottemperanza al D.Lgs. 155/2010 e ss.mm.ii. sono state calcolate le soglie di valutazione superiore ed inferiore per il PM10 e per il PM2.5.

Dalle tabelle riportate in calce e quelle della sezione «Box-Plot», si evince che si verificano superamenti solo della soglia di valutazione inferiore per il PM2.5 nella stazione di Masseria De Blasiis e la soglia di valutazione inferiore per il limite giornaliero del PM10 su tutte le stazioni considerate.

Media giornaliera del PM10 Superamenti della Soglia di valutazione superiore				
	Costa Molina Sud 1	Masseria De Blasiis	Grumento 3	Viggiano1
03/13 - 12/13	5	6	4	3
2014	6	6	8	6
2015	12	13	13	10
2016	20	13	19	20
2017	17	14	8	6
01/18 - 02/18	1	1	1	1

Media giornaliera del PM10 Superamenti della Soglia di valutazione inferiore				
	Costa Molina Sud 1	Masseria De Blasiis	Grumento 3	Viggiano1
03/13 - 12/13	33	63	24	17
2014	33	34	26	25
2015	35	56	40	40
2016	41	43	37	42
2017	40	46	34	26
01/18 - 02/18	6	8	8	1

Conclusioni

Dall'applicazione dei diversi metodi statistici si desumono informazioni analoghe, per cui le quattro stazioni esaminate forniscono dati ridondanti sia per il PM10 sia per il PM2.5. In particolare con il metodo della PCA risulta che le stazioni hanno tutte lo stesso peso per la misura delle concentrazioni. La metodologia descritta sarà estesa all'analisi dei parametri più significativi di qualità dell'aria e meteo rilevati ed a tutte le centraline della rete regionale, integrandola con l'elaborazione dei *polar cluster* tra concentrazione e componenti del vento al fine di individuare eventuali sorgenti emmissive dominanti. Dall'ulteriore verifica del rispetto dei criteri normativi delle Soglie di Valutazione Inferiore e Soglie di Valutazione Superiore emerge che per le misure di PM2.5 nei siti di Viggiano 1, Costa Molina Sud 1 e Grumento 3 si possono utilizzare in via esclusiva le tecniche di modellizzazione o di stima obiettiva; in tutte le altre stazioni considerate le misurazioni in siti fissi, sia di PM10 sia di PM2.5, possono essere combinate con misurazioni indicative o tecniche di modellizzazione (D.Lgs. 155/2010, Art.2 , c.1, lettere z) e aa)).

Bibliografia:

Decreto Legislativo n. 155/2010 e ss.mm.ii.;
Carslaw, D.C. (2015). The openair manual — open-source tools for analyzing air pollution data. Manual for version 1.1-4, King's College London.
On the performance of Fisher Information Measure and Shannon entropy estimators L Telesca, M Lovallo Physica A: Statistical Mechanics and its Applications 484, 569-576;
Comparisons of urban and rural PM10 and PM2.5 mass concentrations and semi-volatile fractions in northeastern Colorado, Nicholas Clements1, Michael P. Hannigan1, Shelly L. Miller, Jennifer L. Peel, and Jana B. Milford;
Statistical Methods in Assessing Agreement: Models, Issues, and Tools, Lawrence Lin, A. S. Hedayat, Bikas Sinha, and Min Yang;
Tukey J.W., Exploratory Data Analysis Addison-Wesley, Reading, Massachusetts, USA, 1977.