



UNIVERSITÀ POLITECNICA DELLE MARCHE

FACOLTÀ DI INGEGNERIA

Corso di Laurea triennale in **INGEGNERIA MECCANICA**

**ANALISI DELLA QUALITÀ DELL'ARIA ATTRAVERSO
L'UTILIZZO DI DISPOSITIVI INDOSSABILI**

AIR QUALITY ANALYSIS USING WEARABLE DEVICES

Relatore:

Prof. Gian Marco Revel

Tesi di Laurea di:

Francesco Marconi

Correlatore:

Dott. Ing. Nicole Morresi

A.A. 2023/2024

Indice

Capitolo 1: Introduzione.....	6
1.1 Contesto e Scopo del Lavoro	6
1.2 Progetto SOCIO-BEE.....	8
1.3 Struttura della tesi	9
Capitolo 2: La misura della qualità dell'aria.....	10
2.1 Inquinanti misurati e rischi correlati.....	10
2.1.1 Materiale particolato aerodisperso (PM2.5)	11
2.1.2 Ozono (O ₃)	12
2.1.3 Diossido di azoto (NO ₂).....	13
2.2 Normative vigenti	13
2.3 Metodi di misura della qualità dell'aria	15
2.3.1 Stazioni di riferimento: metodo gravimetrico	16
2.3.2 Sensori a basso costo: metodo ottico.....	17
2.3.3 Sensori a basso costo: metodo elettrochimico.....	18
Capitolo 3: Lo stato dell'arte.....	19
3.1 La Citizen Science per influenzare il <i>policy-making</i>	19
3.2 La Citizen Science per la misura della qualità dell'aria	20
Capitolo 4: Materiali e metodi.....	26
4.1 Campagne di misura	26
4.1.1 Set-up sperimentale	27
4.1.1.1 La piattaforma SOCIO-BEE	28
4.1.1.2 L'applicazione SOCIO-BEE	30
4.1.1.3 Wearable Sensor Node (WSN).....	31
4.1.2 Ancona.....	35
4.1.3 Maroussi	37
4.1.4 Saragozza	38
4.2 Analisi dati	39
4.2.1 Dataset e pre-processamento dei dati	40
4.2.2 Filtraggio preliminare dei dataset e analisi statistica	41
4.2.3 Ancona: analisi dati	42
4.2.3.1 Analisi di incertezza	43
4.2.3.2 Analisi geografica.....	44
4.2.3.3 Analisi dati delle campagne con drone.....	46
4.2.4 Analisi dati: Maroussi	46

4.2.4.1 Analisi geografica.....	47
4.2.5 Analisi dati: Saragozza	48
Capitolo 5: Risultati	49
5.1 Ancona.....	49
5.1.1 Risultati dell'analisi di incertezza	49
5.1.2 Risultati dell'analisi geografica.....	53
5.1.3 Risultati dell'analisi sulle campagne di misura con drone	54
5.2 Maroussi	55
5.3 Saragozza	57
Capitolo 6: Conclusioni.....	58
Riferimenti bibliografici	60

Indice delle figure

Figura 1: Metafora dell'alveare - Progetto SOCIO-BEE	8
Figura 2: Numero di pubblicazioni sul tema ambientale, per anno di pubblicazione	10
Figura 3: Morti premature per anno in Europa	12
Figura 4: Procedimento del metodo di misura gravimetrico.....	16
Figura 5: Pesatura del filtro nel metodo gravimetrico.....	17
Figura 6: Schema di funzionamento di un sensore ottico	18
Figura 7: Schema di funzionamento di un sensore elettrochimico	18
Figura 8: Grafico sull'occorrenza cumulativa delle keywords per anno di pubblicazione	19
Figura 9: Evidence Co-creation Framework (ECF).....	20
Figura 10: ECF all'interno del processo di policy-making	20
Figura 11: Ubicazione dei sensori OPC-N2 attorno a ML#12	21
Figura 12: Boxplot rappresentanti le distribuzioni di PM2.5 e PM10 nei 15 nuclei	22
Figura 13: a) Area geografica oggetto di studio; b) Stazione di riferimento di Pirkkola; c) Sensore Sensirion SPS30	23
Figura 14: Boxplot PM2.5 e PM10; confronto tra LCS e stazione di riferimento.....	23
Figura 15: Boxplot dei trend di PM2.5 e PM10 indoor e outdoor	24
Figura 16: Heatmap PM2.5 e PM10	24
Figura 17: Fasi del procedimento dello studio.	26
Figura 18: Fasi del procedimento del progetto nelle città pilota.....	27
Figura 19: Piattaforma SOCIO-BEE; campagne attive.....	29
Figura 20: Piattaforma SOCIO-BEE; panoramica campagne	29
Figura 21: Piattaforma SOCIO-BEE; adesione a campagna.....	30
Figura 22: Piattaforma SOCIO-BEE; materiale formativo.....	30
Figura 23: Applicazione SOCIO-BEE; dashboard e informazioni riguardo gli inquinanti	31
Figura 24: Applicazione SOCIO-BEE; campagne proposte	31
Figura 25: Main board WSN	32
Figura 26: Grafico della potenza di trasmissione in funzione della distanza; Bluetooth.....	33
Figura 27: Grafico della potenza di trasmissione in funzione della distanza; Wi-Fi	33
Figura 28: Dimensioni e render del WSN	35
Figura 29: Ubicazione delle misure delle campagne da terra e dei sensori fissi	36
Figura 30: Ubicazione misure con drone	36
Figura 31: Ubicazione stazioni di riferimento.....	37
Figura 32: Ubicazione delle misure delle campagne da terra	38
Figura 33: Ubicazione delle misure delle campagne da terra	39
Figura 34: Dataframe delle campagne da terra.....	41
Figura 35: Dataframe dei sensori fissi	41
Figura 36: Dataframe delle campagne con drone	41
Figura 37: Grafici di confronto tra sensori fissi e stazioni di riferimento.....	49
Figura 38: Grafici dell'analisi di regressione tra sensori fissi e riferimento di Cittadella; pre e post filtraggio	50
Figura 39: Grafici del confronto tra campagne di misura da terra e stazioni di riferimento; pre e post filtraggio	51
Figura 40: Grafici dell'analisi di regressione tra campagne da terra e riferimento di Cittadella; pre e post filtraggio	52
Figura 41: Divisione in cluster del dataset delle campagne da terra.....	53
Figura 42: Confronto temporale tra campagne con drone e campagne ausiliarie da terra	54

Figura 43: Grafici delle misure con drone in funzione dell'indice temporale e dell'altezza sul livello del mare	55
Figura 44: Divisione in cluster dei giorni selezionati	55
Figura 45: Grafico delle misure delle campagne da terra; pre e post filtraggio	57

Indice delle tabelle

Tabella 1: Valori limite per la concentrazione di O ₃ ; D.Lgs. 155/2010	13
Tabella 2: Valori limite per la concentrazione di PM _{2.5} ; D.Lgs. 155/2010	14
Tabella 3: Valori limite per la concentrazione di NO ₂ ; D.Lgs. 155/2010	14
Tabella 4: Valori limite per l'OMS per la concentrazione di PM _{2.5} , O ₃ e NO ₂	15
Tabella 5: Valori di PM _{2.5} e PM ₁₀ nei 15 nuclei	21
Tabella 6: Esposizione personale al PM _{2.5}	25
Tabella 7: Componenti del sensore proposti	32
Tabella 8: Specifiche della batteria VARTA EasyPack	33
Tabella 9: Specifiche del sensore PM _{2.5}	34
Tabella 10: Specifiche dei sensori NO ₂ e O ₃	34
Tabella 11: Analisi riassuntiva delle campagne di Ancona	35
Tabella 12: Analisi riassuntiva delle campagne di Maroussi	37
Tabella 13: Analisi riassuntiva delle campagne di Saragozza	38
Tabella 14: Panoramica sui dati delle tre città pilota	39
Tabella 15: Panoramica sulle possibili analisi per le tre città pilota	39
Tabella 16: Risultati dell'analisi di incertezza; confronto tra campagne di misura da terra e stazione di riferimento	50
Tabella 17: Risultati dell'analisi di incertezza; confronto tra sensore fisso e stazione di riferimento ...	52
Tabella 18: Media e deviazione standard dei cluster e dei riferimenti	53
Tabella 19: Risultati analisi geografica di Ancona	54
Tabella 20: Media e deviazione standard dei cluster dei 10 giorni con più misure	56
Tabella 21: Risultati analisi geografica di Maroussi	56
Tabella 22: Media e deviazione standard delle campagne da terra di Saragozza	57

Capitolo 1: Introduzione

1.1 Contesto e Scopo del Lavoro

In tutto il mondo, uno dei problemi più grandi e impegnativi è il degrado della qualità dell'aria, causato principalmente dalle attività umane che hanno avuto effetti sul cambiamento climatico. Nel 2019, secondo l'Organizzazione Mondiale della Sanità (OMS), il 99% della popolazione mondiale viveva in aree in cui i livelli guida di qualità dell'aria indicati dall'OMS non erano stati raggiunti [1]. Inoltre, l'OMS stima che 4,2 milioni di persone muoiano ogni anno a causa dell'esposizione agli inquinanti atmosferici. Questa preoccupazione è particolarmente evidente nelle aree urbane ad alta densità di popolazione, poiché i livelli di inquinamento atmosferico in tali zone raggiungono spesso valori elevati e, talvolta, allarmanti. Infatti, nonostante le città occupino solo il 2% della superficie terrestre, esse ospitano oltre il 50% della popolazione mondiale [2]. Un'altra questione legata all'inquinamento atmosferico è il significativo costo economico a livello globale: si stima che il mondo spenda il 2-5% del PIL complessivo per curare le malattie correlate all'inquinamento [3].

Data la dimensione del problema, le autorità regolatrici come le Agenzie per la protezione dell'ambiente (EPA) e i responsabili politici di molti Paesi hanno risposto in modo appropriato istituendo ampie reti di monitoraggio della qualità dell'aria e politiche per ridurre l'esposizione personale all'inquinamento atmosferico. Queste azioni si concentrano principalmente sull'utilizzo di sensori di monitoraggio della qualità dell'aria, sull'acquisizione di dati per la formulazione e l'attuazione delle politiche e sulla promozione di cambiamenti comportamentali [4].

Tradizionalmente, le concentrazioni di inquinanti nell'aria sono monitorate utilizzando stazioni di monitoraggio della qualità dell'aria professionali che soddisfano criteri di misura rigorosi e normati. Queste stazioni sono molto accurate ma anche estremamente costose, con costi che possono arrivare a centinaia di migliaia di dollari per singola unità. Anche il loro funzionamento richiede una manutenzione periodica da parte di tecnici specializzati, il che contribuisce a limitarne la diffusione: molte aree metropolitane dispongono di una sola stazione [5]. Inoltre, come sottolineato nello studio [6], un ulteriore limite delle stazioni professionali è la frequenza con cui i dati sono resi disponibili, spesso oraria o persino giornaliera, limitando così la copertura temporale e riducendo la precisione dell'analisi. Questo causa una forte limitazione nella risoluzione temporale del monitoraggio, non permettendo perciò un'analisi accurata e provocando una perdita di generalità delle informazioni derivanti da essa.

Nonostante l'OMS raccomandi una stazione di monitoraggio della qualità dell'aria per chilometro quadrato, e la direttiva UE sull'aria pulita suggerisca una stazione ogni circa 200.000 abitanti [5], un monitoraggio efficace dell'inquinamento atmosferico richiederebbe l'installazione di centinaia di stazioni in aree geografiche ristrette, permettendo così una maggiore densità spaziale di misure.

In questo contesto, gli studi su reti di sensori a basso costo (Low-Cost Sensors, LCS) rivestono un ruolo molto importante, con particolare attenzione all'organizzazione di queste reti. Infatti, come sostiene lo studio [7], recenti progressi tecnologici e la proliferazione dei sensori di qualità dell'aria offrono ulteriori possibilità di affinare la caratterizzazione spazio-temporale dei livelli di inquinamento atmosferico. Oggi sono disponibili numerosi strumenti sviluppati da aziende commerciali, organizzazioni non profit e startup. Sebbene le prestazioni di questi sensori possano variare considerevolmente tra i diversi modelli, essi offrono la possibilità di generare un grande volume di dati a costi ridotti.

In questo ambito, la Citizen Science (CS) gioca un ruolo fondamentale. La Citizen Science è un processo in cui comunità e individui sono coinvolti nella progettazione e nell'esecuzione di esperimenti scientifici con un coinvolgimento di scienziati professionisti [8]. Perciò, nella maggior parte degli studi che coinvolgono i cittadini, il monitoraggio ambientale viene eseguito utilizzando sensori a basso costo distribuiti agli utenti, ai quali viene chiesto di svolgere determinati compiti con lo scopo di ottenere un grande numero di dati in un relativamente breve lasso di tempo, andando così a migliorare nettamente la copertura temporale e cercando anche di coinvolgere e sensibilizzare i cittadini sul problema della qualità dell'aria e come le attività umane possono fortemente influenzarla.

Da questi studi emerge un importante quesito riguardante la qualità dei dati raccolti dai LCS in quanto questi sensori offrono un'incertezza superiore rispetto alle stazioni di monitoraggio ufficiali. Infatti, potendo i sensori essere soggetti ad anomalie e influenze esterne – in particolare umidità relativa e temperatura –, non è possibile completare una calibrazione completa dei sensori a basso costo, e perciò si rende necessario svolgere molto lavoro per utilizzare queste raccolte di dati e valutarne la qualità. Tuttavia, quando vengono utilizzate informazioni contestuali come meteo, posizione o tempo di misura per l'elaborazione dei dati, i LCS portano vantaggi significativi nell'ambito del monitoraggio [6].

Date queste premesse, questo lavoro di tesi si pone un duplice obiettivo. Per prima cosa, lo studio, che si inquadra all'interno di un progetto europeo Citizen Science ha posto la sua attenzione sulla qualità dei dati ottenuti con i LCS, andando ad analizzare il dataset fornito per comprendere l'incertezza di misura dei sensori utilizzati e l'impatto che la non professionalità dell'utente può aver avuto sulla misura stessa. L'impostazione data al nostro studio ha seguito perciò questa direzione, andando a coinvolgere un numero di utenti maggiore rispetto ai precedenti studi per avere così un campione più ampio ed eseguendo campagne di misura in diverse città di diversi Paesi; e in seconda battuta grazie a un'analisi dati di incertezza – riprendendo le modalità evidenziate ad esempio nella review [9] – è stato possibile comprendere come rendere utilizzabili i dati in contesti scientifico-politici.

In secondo luogo, si è indagato sull'influenza dell'ubicazione e della distanza spaziale tra le misure, per dimostrare la necessità di una rete distribuita e riflettere sull'effettiva densità di

tale rete. Gli studi [7] e [10] evidenziano l'importanza e il valore aggiunto della rete di monitoraggio indipendente in termini di disponibilità dei dati, copertura e accesso in tempo reale agli open data, dimostrando inoltre il potenziale della Citizen Science di essere una risorsa per il *policy-making*. In quest'ottica, il nostro studio va a dimostrare – attraverso un'analisi geografica dei dati raccolti – l'effettiva dipendenza che sussiste tra coordinata GPS e misura.

1.2 Progetto SOCIO-BEE

Questo studio fa parte del progetto SOCIO-BEE (Wearables and droneS fOr City socio-environmental Observations and BEhavioral changE, <https://socio-bee.eu/>), un'iniziativa finanziata dal programma dell'Unione Europea "Horizon 2020 Framework Programme for Research and Innovation". Il progetto – sviluppato in tre città europee: Ancona, Italia; Maroussi, Grecia; Saragozza, Spagna – pone le sue fondamenta sulla Citizen Science, sfruttata per delineare ed eseguire esperimenti con l'ambiente come argomento centrale, affiancato da temi quali l'innovazione e la sostenibilità con un focus particolare sulla qualità dell'aria. Gli scopi del progetto si possono ricondurre a due generali tematiche: la prima, aumentare la consapevolezza e incoraggiare le comunità a monitorare l'ambiente, permettendogli di venire a conoscenza di comportamenti sostenibili da tenere; la seconda, permettere alle suddette comunità di influenzare le future policy riguardanti il tema ambientale.

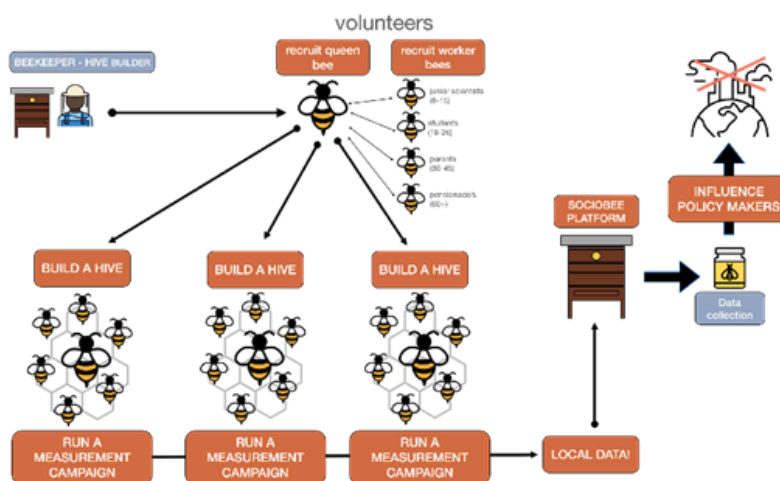


Figura 1: Metafora dell'alveare - Progetto SOCIO-BEE

Le modalità con cui il progetto mira a raggiungere tali scopi si delineano attorno alla partecipazione attiva del cittadino nella misura della qualità dell'aria, attraverso l'utilizzo di sensori a basso costo indossabili. Tali utenti sono supportati da funzionari locali, ricercatori e gruppi di azione. La metafora utilizzata per descrivere i compiti di ciascun partecipante è quella dell'alveare, come mostrata in figura [1].

Così come nell'alveare, il contributo e il lavoro di tutti i partecipanti è importante e necessario:

- Le api operaie effettuano misure per raccogliere dati in diverse aree utilizzando un cellulare e un sensore.
- Le api regine progettano le campagne a cui puoi partecipare.
- Gli apicoltori si prenderanno cura dell'alveare e si assicureranno che cresca.
- Gli orsi analizzano i dati raccolti con i sensori e il cellulare per fare politiche migliori.

Ottenute le misure della qualità dell'aria, queste vengono analizzate, discusse e si traggono conclusioni sulle misure. Queste conclusioni ci aiuteranno a convalidare ipotesi volte a mitigare e ridurre l'inquinamento atmosferico e ci permetteranno di influenzare i processi decisionali per ridurre i livelli di inquinamento nelle città e creare fiducia tra cittadini e istituzioni.

1.3 Struttura della tesi

Nei seguenti paragrafi verrà innanzitutto approfondito il tema centrale dell'analisi, ovvero la misura della qualità dell'aria, introducendo gli inquinanti usualmente monitorati, i rischi per la salute a loro annessi e le normative vigenti per il loro controllo; si svolge poi una breve analisi dei metodi di misura di tali inquinanti, trattando delle stazioni di monitoraggio e dei sensori a basso costo, e delle tecniche di misure utilizzate dai sensori. Si presentano poi gli studi correlati, con i temi di Citizen Science, LCS e analisi dati della misura della qualità dell'aria centrali alle loro analisi. Si prosegue con la descrizione dei materiali utilizzati per il nostro studio – in particolare esponendo le metodologie di organizzazione e misura utilizzate nelle tre città pilota del progetto SOCIO-BEE –, la struttura del programma Python sviluppato per elaborare i dati raccolti e il procedimento di analisi dei suddetti dati. Infine, si discutono i risultati, le conclusioni, i limiti dell'esperimento e vengono inoltre presentati alcuni spunti per futuri studi.

Capitolo 2: La misura della qualità dell'aria

La qualità dell'aria è un indicatore fondamentale dello stato di salute ambientale, perciò il suo monitoraggio ha assunto una sempre maggiore importanza, specialmente negli ultimi decenni – come dimostra la figura [2], nella quale è evidenziato come il numero di pubblicazioni riguardanti l'argomento è aumentato anno per anno [5] –, nei quali con l'evoluzione della nostra società e l'aumento delle fonti di inquinamento nell'ambiente, l'impatto dell'inquinamento atmosferico è diventato sempre più evidente [11].

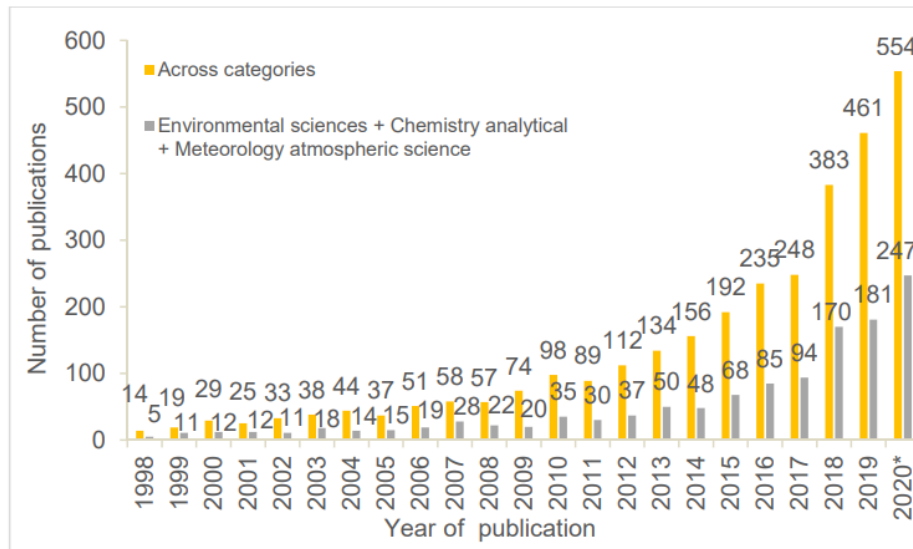


Figura 2: Numero di pubblicazioni sul tema ambientale, per anno di pubblicazione

L'aria che respiriamo contiene una varietà di sostanze, alcune delle quali – che vanno sotto il nome di inquinanti – possono essere nocive per l'uomo e per l'ambiente. Tali inquinanti, che comprendono ad esempio il particolato fine (PM10 e PM2.5), gli ossidi di azoto (NO_x), l'ozono (O₃), il biossido di zolfo (SO₂), il monossido di carbonio (CO) e i composti organici volatili (COV), influiscono sulla salute fisica e mentale delle persone e l'esposizione a lungo termine aumenta il rischio di malattie cardiovascolari e respiratorie [2].

Il monitoraggio di questi inquinanti può essere svolto in varie modalità, con l'utilizzo di stazioni di riferimento ufficiali o con l'uso di sensori a basso costo. Entrambe le soluzioni sono valide alternative per raccogliere dati essenziali per l'analisi e la gestione della qualità dell'aria, per poter fornire una migliore informazione alle autorità permettendogli così di formulare normative – internazionali e locali – per limitare la concentrazione di tali inquinanti.

2.1 Inquinanti misurati e rischi correlati

In questo studio e come si può notare dai lavori citati fino ad ora, l'attenzione è posta principalmente su tre di tutti gli inquinanti sopra menzionati: il materiale particolato

aerodisperso (più precisamente il PM2.5), l'ozono (O₃) e il diossido di azoto (NO₂). Si propone ora una panoramica su questi tre inquinanti, andandoli prima di tutto a definire, per poi presentare i rischi ad essi annessi.

2.1.1 Materiale particolato aerodisperso (PM2.5)

Come si può leggere nel rapporto *Concentrazione atmosferica degli inquinanti* del Ministero dell'Ambiente e della Tutela del Territorio e del Mare [13] “per materiale particolato aerodisperso si intende l'insieme delle particelle atmosferiche solide e liquide sospese in aria ambiente. Il termine PM2.5 identifica le particelle di diametro aerodinamico inferiore o uguale ai 2,5 [µm] [...]. Il particolato PM2.5 è detto anche ‘particolato fine’, denominazione contrapposta a ‘particolato grossolano’ che indica tutte quelle particelle sospese con diametro aerodinamico maggiore di 2,5 [µm] [...]. Le sorgenti del particolato fine sono un generalmente tutti i tipi di combustione, inclusi quelli dei motori di auto e motoveicoli, degli impianti per la produzione di energia, della legna per il riscaldamento domestico, degli incendi boschivi e di molti altri processi industriali.”

Il PM2.5 è una delle sostanze inquinanti più dannose per la salute, in quanto – come si può leggere da [12] – “data la sua dimensione estremamente ridotta, è capace di penetrare in profondità nel sistema respiratorio umano, raggiungendo non solo la trachea e le vie respiratorie superiori, quale è il caso del PM10, ma anche gli alveoli polmonari. Il particolato fine ha effetti nocivi sia sul sistema respiratorio che su quello circolatorio. Secondo le analisi dell'OMS, una esposizione prolungata ha comprovati legami con l'emergere di tumori e di altre patologie come l'obesità, il diabete, ma anche il morbo di Alzheimer e la demenza. Può inoltre causare arteriosclerosi e, secondo ricerche recenti, potrebbe incidere sullo sviluppo neurologico nei bambini e sulle funzioni cognitive negli adulti. Oltre a esacerbare problemi di salute preesistenti.”

Nella figura [3] è riportato un grafico – estratto anch'esso da [12] – rappresentante le morti premature in Europa legate alla presenza del particolato fine. Si sottolinea come per ‘morti premature’ si intenda il decesso anticipato rispetto all'aspettativa di vita nei Paesi in esame.

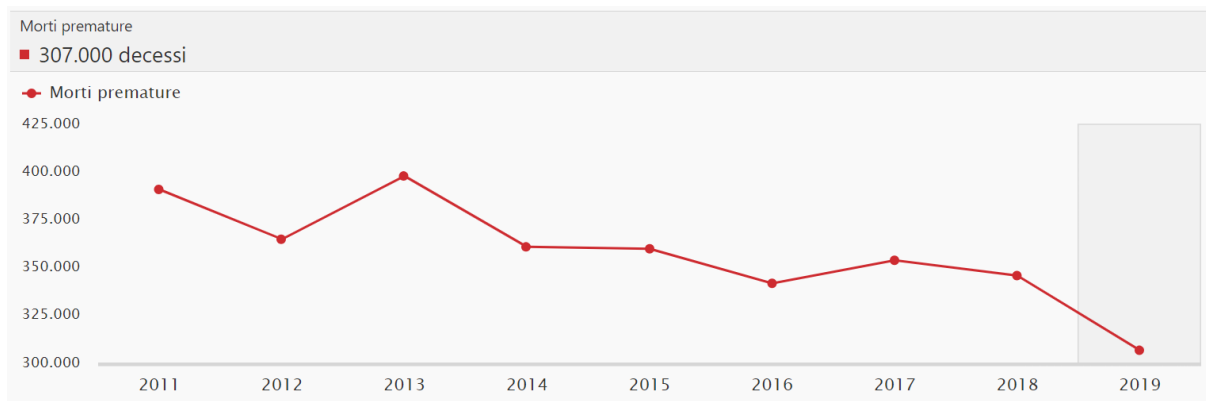


Figura 3: Morti premature per anno in Europa

2.1.2 Ozono (O₃)

Riprendendo un opuscolo pubblicato dal Ministero della Salute nel 2015 [14], “l’ozono è un gas composto da molecole instabili con un odore pungente e dotato di grande reattività. Viene prodotto in atmosfera dalla reazione tra ossidi di azoto, composti organici volatili e raggi solari. In genere, la quota proveniente dall’esterno rappresenta la maggior parte dell’ozono presente in un ambiente confinato, tuttavia, nelle abitazioni può essere emesso in maniera significativa da strumenti elettrici ad alto voltaggio, quali motori elettrici, stampanti laser e fax, da apparecchi che producono raggi ultravioletti, da filtri elettronici per pulire l’aria, non correttamente installati e senza una adeguata manutenzione.”

Assieme al PM2.5, l’ozono è uno degli inquinanti più nocivi per l’essere umano. Come riportato in un rapporto dell’Istituto Superiore di Sanità [15] “si stima che circa 21 mila decessi prematuri all’anno siano associati a livelli di ozono che superano i 70 [µg/m³] nell’Europa a 25 Paesi. [...] All’ozono sono associati anche i 14 mila ricoveri ospedalieri per malattie respiratorie, che si registrano ogni anno nei 25 Paesi dell’Unione europea.” Infatti, citando nuovamente l’opuscolo [14], i rischi legati all’esposizione all’ozono sono “effetti irritativi alle mucose oculari e alle prime vie aeree, tosse, fenomeni broncostruttivi ed alterazione della funzionalità respiratoria. In studi epidemiologici condotti in popolazioni urbane esposte ad ozono sono stati osservati sintomi irritativi sulle mucose oculari e sulle prime vie respiratorie per esposizioni di alcune ore a livelli di ozono a partire da 0,2 mg/m³ (media oraria). In bambini ed in giovani adulti sono state osservate riduzioni transitorie della funzionalità respiratoria, a livelli inferiori di ozono, a partire da 0,12 mg/m³ (media oraria). Sono invece disponibili pochi studi sugli effetti per esposizioni croniche a questo inquinante.”

2.1.3 Diossido di azoto (NO₂)

Riprendendo il rapporto *Concentrazione atmosferica degli inquinanti* del Ministero dell'Ambiente e della Tutela del Territorio e del Mare [16], "il biossido di azoto (NO₂) è un gas di colore bruno-rossastro, poco solubile in acqua, tossico, dall'odore forte e pungente e con forte potere irritante. È un inquinante a prevalente componente secondaria, in quanto è il prodotto dell'ossidazione del monossido di azoto (NO) in atmosfera; solo in proporzione minore viene emesso direttamente in atmosfera. La principale fonte di emissione degli ossidi di azoto (NO_x=NO+NO₂) è il traffico veicolare; altre fonti sono gli impianti di riscaldamento civili e industriali, le centrali per la produzione di energia e un ampio spettro di processi industriali."

Anche il diossido di azoto, assieme ai due inquinanti sopra citati, è estremamente nocivo. Infatti, come si può leggere in [17], "si tratta di una sostanza particolarmente nociva sia per l'ambiente che per la nostra salute. A lungo andare, contribuisce all'eutrofizzazione, ovvero un processo degenerativo degli ecosistemi acquatici, dovuto a un eccesso di sostanze nutritive, e all'acidificazione delle piogge – due importanti e nocivi effetti del cambiamento climatico. Da un punto di vista sanitario invece un'esposizione prolungata contribuisce all'emergere di gravi patologie quali il diabete, l'ipertensione e i problemi cardiovascolari. Oltre a ostacolare il normale sviluppo della capacità polmonare. Studi recenti hanno inoltre rilevato un più elevato tasso di mortalità negli individui che abitano in zone maggiormente esposte a questa sostanza."

2.2 Normative vigenti

Le concentrazioni degli inquinanti appena descritti sono regolamentate in Italia dal D.Lgs. 155/2010, che si pone come obiettivo quello di consentire, a regioni e province autonome, la valutazione e la gestione della qualità dell'aria ambiente.

Nelle tabelle [1], [2] e [3] riportate di seguito – estratte dal rapporto *Concentrazione atmosferica degli inquinanti* – è possibile visualizzare i valori limite ciascun inquinante menzionato, rispettivamente O₃, PM2.5 e NO₂.

Tabella A: O₃ - Soglia di informazione, soglia di allarme, obiettivo a lungo termine per la protezione della salute umana e per la protezione della vegetazione ai sensi del D.Lgs. 155/2010

	Valore	Periodo di mediazione
Soglia di informazione	180 µg/m ³	1 ora
Soglia di allarme	240 µg/m ³	1 ora
Obiettivo a lungo termine per la protezione della salute umana	120 µg/m ³	Media massima giornaliera calcolata su 8 ore
Obiettivo a lungo termine per la protezione della vegetazione (AOT40v)	6.000 µg/m ³ *h	1 ora cumulativa da maggio a luglio

Tabella 1: Valori limite per la concentrazione di O₃; D.Lgs. 155/2010

Tabella A: PM_{2,5} - valore limite per la salute umana

	Periodo di mediazione	Valore limite	Margine di tolleranza	Data alla quale il valore limite deve essere raggiunto
FASE 1				
Valore limite annuale	Anno civile	25 µg/m ³	20 % all'11 giugno 2008, con riduzione il 1 gennaio successivo e successivamente ogni 12 mesi secondo una percentuale annua costante fino a raggiungere lo 0 % entro il 1° gennaio 2015	1° gennaio 2015
FASE 2				
Valore limite annuale	Anno civile	20 µg/m ³		1° gennaio 2020

Tabella B: PM_{2,5} - valore limite di concentrazione aumentato del margine di tolleranza

Periodo di mediazione	Valore limite più margine di tolleranza al 2010
Anno civile	29 µg/m ³

*Tabella 2: Valori limite per la concentrazione di PM_{2,5}; D.Lgs. 155/2010***Tabella A: NO₂ - Valori limite per la protezione della salute umana**

	Periodo di mediazione	Valore limite
Valore limite orario	1 ora	200 µg/m ³ NO ₂ da non superare più di 18 volte per anno civile
Valore limite annuale	Anno civile	40 µg/m ³ NO ₂

Tabella 3: Valori limite per la concentrazione di NO₂; D.Lgs. 155/2010

L'Organizzazione Mondiale della Sanità (OMS), pone invece un valore limite di riferimento ben diverso relativamente ai possibili effetti sulla salute, con la soglia massima posta visualizzabili nella Tabella [4] riportata di seguito – estratta dal documento [18].

Inquinante	Tempo di media	Obiettivo intermedio				Livello AQG
		1	2	3	4	
PM_{2,5}, µg/m³	Annuale	35	25	15	10	5
	24 ore ^a	75	50	37,5	25	15
PM₁₀, µg/m³	Annuale	70	50	30	20	15
	24 ore ^a	150	100	75	50	45
O₃, µg/m³	Picco stagionale ^b	100	70	–	–	60
	8 ore ^a	160	120	–	–	100
NO₂, µg/m³	Annuale	40	30	20	–	10
	24 ore ^a	120	50	–	–	25
SO₂, µg/m³	24 ore ^a	125	50	–	–	40
CO, mg/m³	24 ore ^a	7	–	–	–	4

^a 99° percentile (ovvero 3-4 giorni di superamento all'anno).

^b Media della concentrazione media giornaliera massima su 8 ore di O₃ nei sei mesi consecutivi con la più alta concentrazione media mobile semestrale di O₃.

Tabella 4: Valori limite per l'OMS per la concentrazione di PM_{2.5}, O₃ e NO₂

2.3 Metodi di misura della qualità dell'aria

Gli inquinanti ambientali fin ora descritti sono generalmente monitorati da stazioni fisse di riferimento gestite da enti regionali e/o nazionali per la tutela ambientale. Esse – riassumendo quanto detto nel Capitolo 1 di questo studio – hanno un elevato grado di accuratezza e precisione di misura, però al contempo presentano limitazioni nell'ambito della copertura spaziale e temporale di monitoraggio. Ciò deriva dal fatto che – essendo interessate da elevati costi di approvvigionamento, montaggio e manutenzione – sono disponibili in quantità limitata, risultando così in una bassa densità spaziale e inoltre, dovendo seguire le normative ufficiali, le osservazioni degli inquinanti sono riportate con una bassa frequenza temporale.

Queste limitazioni hanno portato i ricercatori – come evidenziato negli studi citati nel Capitolo 1 – a rivolgere la loro attenzione verso i cosiddetti sensori di misura a basso costo. Essi infatti, avendo un costo che si aggira tra i 100\$ e i 2000\$ – contro un costo di due ordini di grandezza maggiore per le stazioni di riferimento – [19], permettono un utilizzo su più grande scala, consentendo perciò una distribuzione spaziale più capillare e una maggiore frequenza di misura.

I due diversi sistemi di monitoraggio appena elencati presentano un'ulteriore differenza fondamentale, che risiede nel metodo di misura utilizzato dal sensore.

Si svolge ora una rassegna dei metodi di misura maggiormente utilizzati, ponendo l'attenzione sul monitoraggio del particolato fine (PM_{2.5}).

2.3.1 Stazioni di riferimento: metodo gravimetrico

La direttiva che regola le stazioni di riferimento è la norma UNI EN 12341:2023, che specifica un metodo di riferimento per la determinazione della concentrazione in massa del particolato sospeso (PM10 o PM2.5) in aria ambiente, attraverso la filtrazione mediante appositi filtri e la pesatura di essi tramite bilancia. Il flusso nominale di campionamento – che avviene con campionatori con uno specifico ugello – è di 2,3 [m³/h], per un periodo nominale di 24 ore. I risultati delle misure sono espressi in [μg/m³]. L'applicabilità di questa norma va da circa 1 [μg/m³], che è il limite di rilevabilità del metodo – espresso come sua incertezza – fino a 120 [mg/m³] per il PM2.5.

Il procedimento completo di misura – rappresentato in figura [4] – è il seguente:

- I filtri puliti vengono pesati in laboratorio
- I filtri vengono caricati nelle apposite cartucce etichettate, preconditionati e trasportati alla stazione di riferimento
- Presso la stazione di riferimento, l'aria o i gas vengono fatti passare negli ugelli prima menzionati con il flusso nominale per il periodo nominale. Il PM2.5 presente viene trattenuto nel filtro
- I filtri carichi vengono nuovamente trasportati in laboratorio, ricondizionati e pesati
- La determinazione della massa di PM2.5 avviene per sottrazione del peso del filtro carico da quello del filtro pulito (pesata differenziale). La concentrazione di PM2.5 viene calcolata utilizzando il flusso nominale e il peso ottenuto.

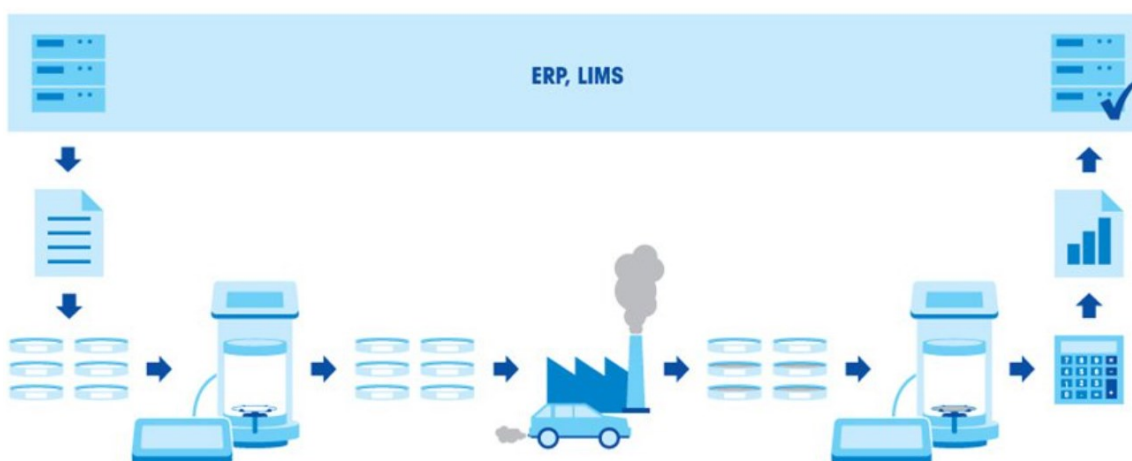


Figura 4: Procedimento del metodo di misura gravimetrico

Essendo i filtri estremamente sensibili alle condizioni termo-igrometriche, la procedura di analisi del peso – eseguita con bilance con tolleranza al [μg], come si può notare in figura [5] –

e il condizionamento dei filtri si effettua in ambienti a condizioni ambientali costanti, come ad esempio una camera bianca:

- Temperatura: $20 \pm 1^\circ\text{C}$
- Umidità: $50 \pm 5\%$



Figura 5: Pesatura del filtro nel metodo gravimetrico

2.3.2 Sensori a basso costo: metodo ottico

Il metodo in questione basa il suo funzionamento sul principio della diffusione ottica (detta anche dispersione o scattering). Un raggio luminoso che procede in una certa direzione e quando esso colpisce una particella, viene diffuso in maniera disordinata in tutte le direzioni.

I sensori sono tipicamente costituiti da una ventola che aspira l'aria esterna per convogliarla in una camera di misura stagna, dove è contenuto un diodo laser e un fotodiodo rivelatore. Una volta immessa l'aria nella camera, il laser viene acceso e il fascio di luce procede senza diffondersi. Quando il laser colpisce il particolato, la sua luce viene diffusa e l'intensità luminosa che si disperde nella camera può essere rilevata dal fotodiodo. Il sensore riesce a distinguere tra il PM10 e il PM2.5 perché la forma d'onda della luce rilevata dal fotodiodo è correlata con il numero e le dimensioni delle particelle [20]. In figura [6], estratta dallo studio [10], è rappresentato lo schema di funzionamento appena descritto.

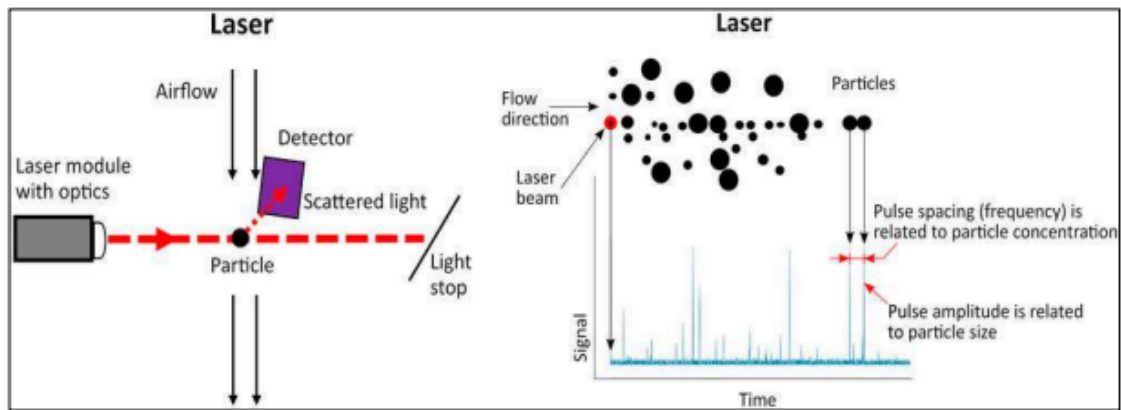


Figura 6: Schema di funzionamento di un sensore ottico

2.3.3 Sensori a basso costo: metodo elettrochimico

I sensori elettrochimici misurano la concentrazione di particolato fine, a differenza di quelli ottici che misurano il numero di particelle. Il loro principio di funzionamento si basa sull'utilizzo di una corrente, che è linearmente proporzionale alla frazione volumetrica del gas da misurare.

Un sensore elettrochimico – il quale schema è rappresentato in figura [7] – è solitamente costituito da due elettrodi e da un elettrolita. La concentrazione del gas viene misurata sull'elettrodo di rilevamento, costituito da un metallo catalitico che ottimizza la reazione del gas da misurare. Il gas reagisce con l'elettrodo di rilevamento dopo essere entrato nel sensore attraverso una barriera di diffusione capillare. Il contro-elettrodo funge da semi-cella e completa il circuito. Il flusso di corrente viene misurato da un circuito esterno, per determinare la concentrazione di gas [21].

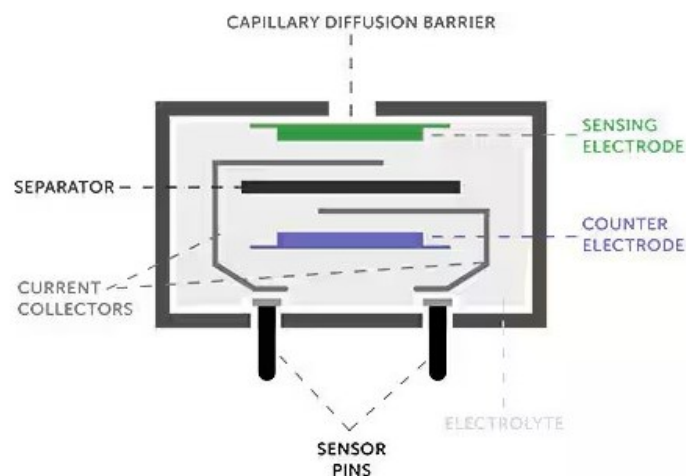


Figura 7: Schema di funzionamento di un sensore elettrochimico

Capitolo 3: Lo stato dell'arte

Negli ultimi anni, numerosi studi hanno sviluppato progetti di Citizen Science con l'obiettivo di valutare l'efficacia nel monitoraggio della qualità dell'aria, approfondendo le opportunità e le limitazioni connesse. Gli approcci adottati in questi studi sono stati molteplici e diversificati.

3.1 La Citizen Science per influenzare il *policy-making*

Sono state svolte review della letteratura per analizzarne le peculiarità e le potenzialità di diventare uno strumento per influenzare (e supportare) il *policy-making* e aumentare la consapevolezza del cittadino sui temi ambientali. Lo studio [4] va prima di tutto a svolgere un'analisi bibliometrica, per comprendere come la Citizen Science sia diventata nell'ultimo periodo un argomento ampiamente discusso – come si può notare dalla Figura [8] nella quale è rappresentato un grafico che mostra l'occorrenza cumulativa delle otto *keywords* più frequenti.

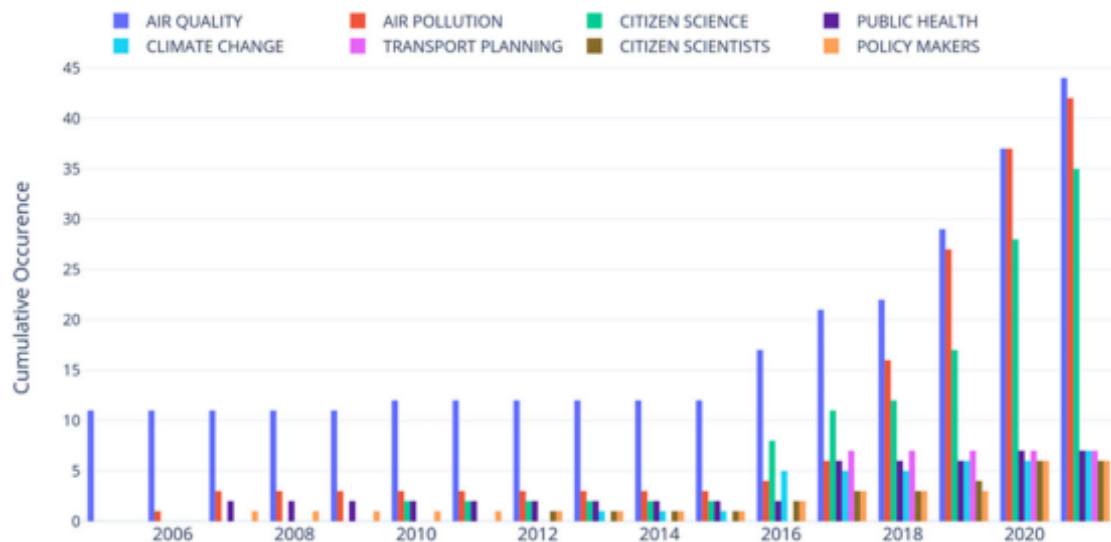


Figura 8: Grafico sull'occorrenza cumulativa delle *keywords* per anno di pubblicazione

L'analisi svolta dallo studio [4] prosegue – saggiando progetti in tutto il mondo – approfondendo gli aspetti della Citizen Science inerenti all'approccio di tali progetti ai dati raccolti e alla possibilità di utilizzare tali dati per rendere il cittadino parte attiva nel processo decisionale delle *policy* ambientali.

Lo studio conclude che la Citizen Science ha le potenzialità per diventare uno strumento del *policy-making*, in quanto è in grado di creare delle valide prove per permettere alla comunità di influenzare la formulazione e implementazione delle *policy* – si guardino le figure [9] e [10], nelle quali sono rappresentati l'*Evidence Co-Creation Framework* (ECF), ovvero le modalità con cui creare le suddette prove, e come l'ECF si possa inserire nel processo decisionale.

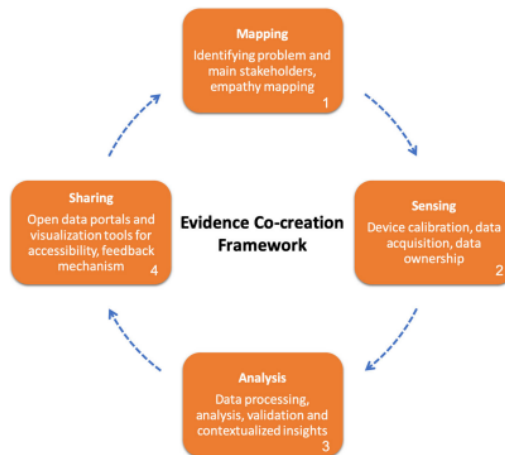


Figura 9: Evidence Co-creation Framework (ECF)



Figura 10: ECF all'interno del processo di policy-making

3.2 La Citizen Science per la misura della qualità dell'aria

Negli ultimi due decenni – come evidenziato fin ora nel nostro studio – il numero di progetti di Citizen Science per la misura della qualità dell'aria è aumentato notevolmente.

Questi studi – ciascuno con i propri scopi e modalità – sono andati ad analizzare differenti aspetti della CS per trarre conclusioni sulle caratteristiche che tali progetti devono possedere per essere di utilità nel monitoraggio della qualità dell'aria.

Nello studio [22], svolto nella capitale del Kenya Nairobi nel 2023, si analizza l'impatto della costruzione di una nuova autostrada – chiamata *Missing Link #12* (ML#12). Il progetto ha coinvolto 15 nuclei familiari che vivono a una distanza massima di 150 metri da ML#12 (nella Figura [11] si possono notare le ubicazioni dei sensori), divisi in sottogruppi da 5 ciascuno, ai quali sono stati distribuiti sensori ottici AlphaSense OPC-N2 dal costo di circa 250 GBP. Le misure sono state svolte in due turni dalla durata di 24 ore ognuno, per un totale di 48 ore.

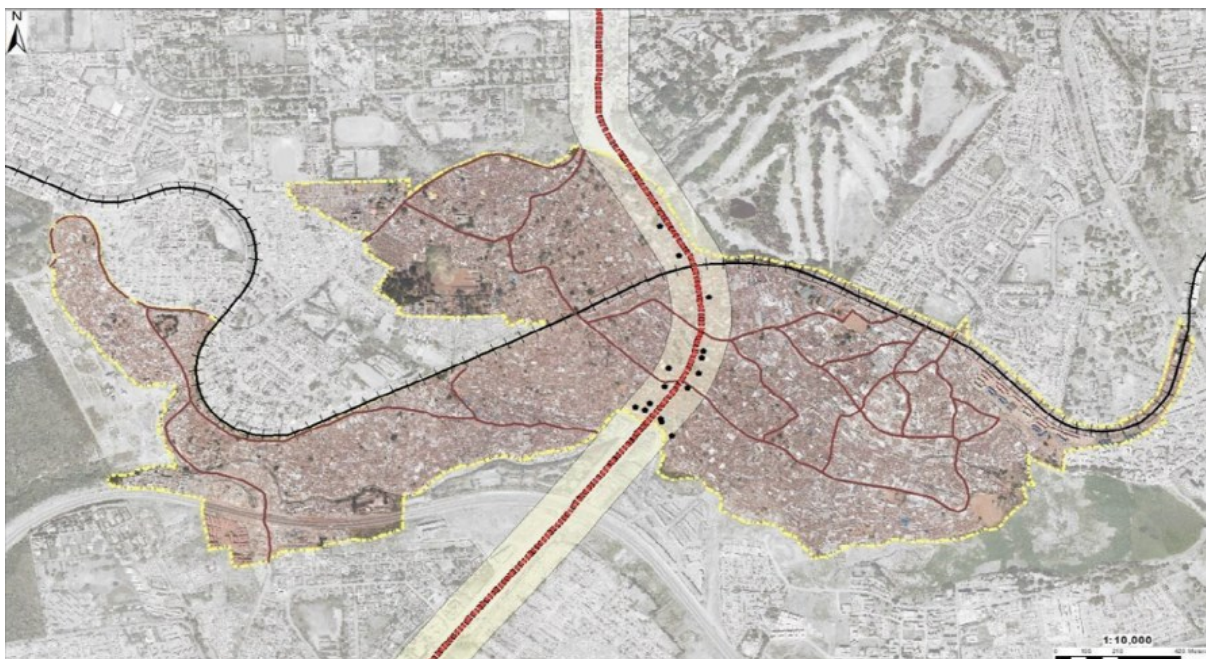


Figura 11: Ubicazione dei sensori OPC-N2 attorno a ML#12

I dati raccolti hanno portato a concludere che la qualità dell'aria nelle zone oggetto delle misure è al di fuori della soglia raccomandata dall'OMS. Più precisamente la media delle misure delle 48 ore nei 15 nuclei familiari è di $28,7 \text{ } [\mu\text{g}/\text{m}^3]$ – dati mostrati nella tabella [6], mentre i valori limiti espressi dall'OMS sono di $5 \text{ } [\mu\text{g}/\text{m}^3]$ per la media annuale e di $15 \text{ } [\mu\text{g}/\text{m}^3]$ per la media giornaliera. Nella figura [12] è presentato un *boxplot* che mostra la distribuzione di PM_{2.5} e PM₁₀ nei 15 nuclei, con una divisione dei dati in base alla fonte dell'inquinante.

Pollutants	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
PM _{2.5}	2.3	14.4	23.1	28.7	34.3	404.5
PM ₁₀	3.5	27.4	43.0	63.5	64.2	1053.8

Tabella 5: Valori di PM_{2.5} e PM₁₀ nei 15 nuclei

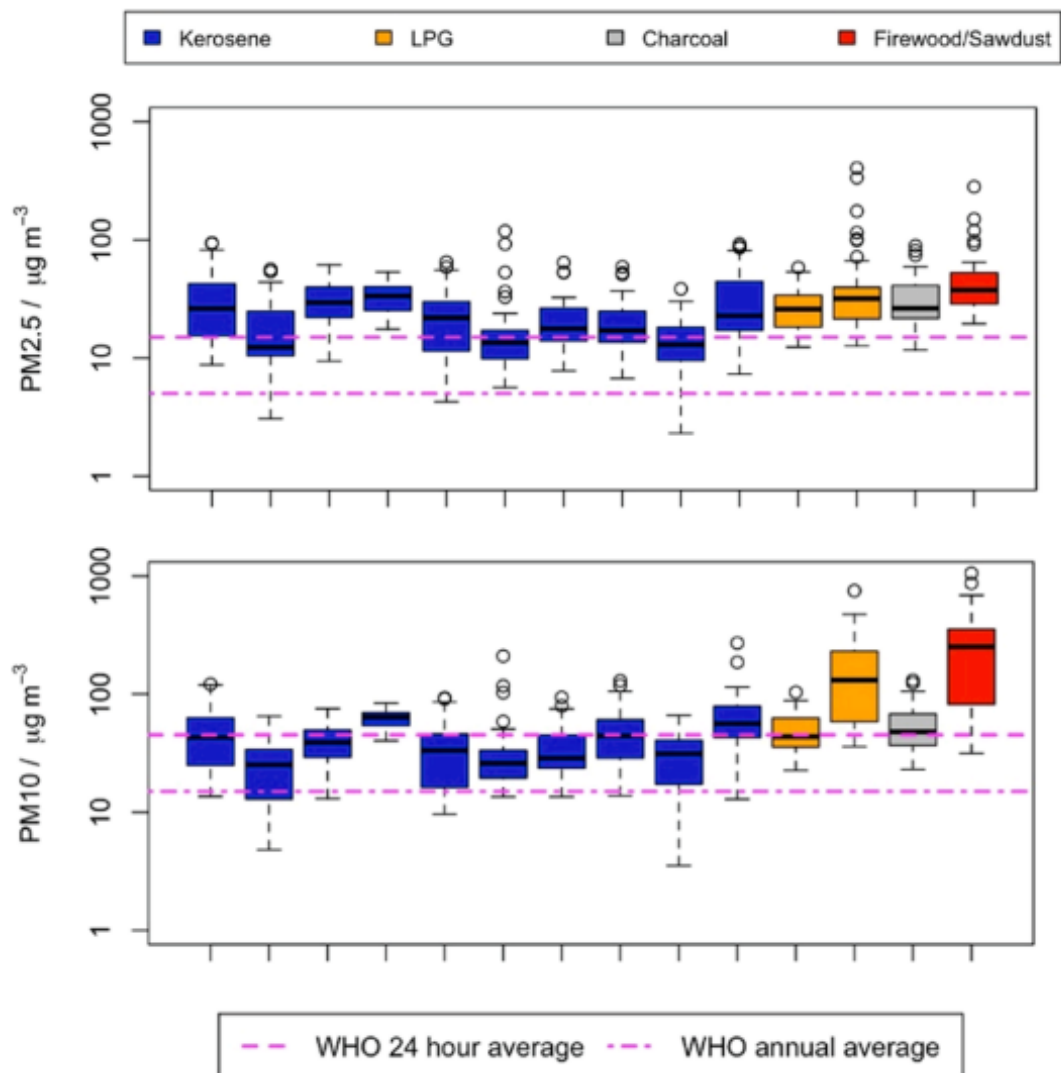


Figura 12: Boxplot rappresentanti le distribuzioni di PM2.5 e PM10 nei 15 nuclei

Nello studio [11], gli scopi sono invece stati:

- Dimostrare l'efficacia della campagna di misura mostrando l'alta risoluzione spaziale del monitoraggio, confrontando i dati raccolti con quelli di una stazione di riferimento.
- Approfondire lo studio degli inquinanti nella zona in esame e valutarne le possibili fonti.
- Verificare le informazioni offerte dal sensore riguardanti l'esposizione personale del cittadino durante il corso della sua giornata.

Le modalità del progetto di CS sono state le seguenti:

- Il sensore Sensirion SPS30 – la cui unità si basa su BMD-340 System On a Module e con una connessione Bluetooth LE a cellulari Android – misurano il particolato (PM), i parametri termo-igrometrici, composti gassosi (e.g. CO₂, NO_x), con un costo di 250\$ in bassi volumi di produzione e meno di 100\$ in alti – si guardi Figura [13.c].

- Un totale di 40 sensori è stato distribuito ai cittadini, ai quali non sono state date indicazioni sui luoghi in cui effettuare misure.
- Sono stati utilizzati per l'analisi solamente i dati in un'area di circa 6 km² – si guardi Figura [13.a] per visualizzare tale area e Figura [13.b] per la stazione.
- La stazione di riferimento utilizzata dallo studio è situata nel quartiere di Pirkkola, all'interno dell'area in esame (si guardi Figura [13.a]).

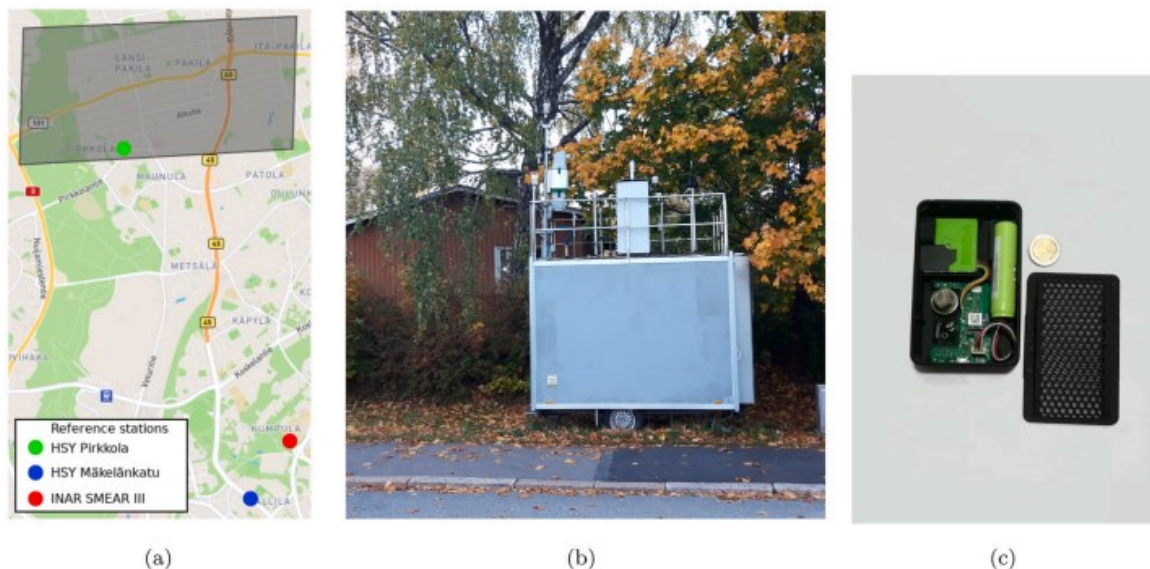


Figura 13: a) Area geografica oggetto di studio; b) Stazione di riferimento di Pirkkola; c) Sensore Sensirion SPS30

I risultati dello studio evidenziano come i valori misurati dai sensori a basso costo non si discostino molto da quelli della stazione di riferimento – come si può vedere nella Figura [14]. Le medie di PM_{2.5} sono infatti rispettivamente 3 [µg/m³] e 4.5 [µg/m³] e i valori di confine del *boxplot* in Figura [14] sono di 1,5 [µg/m³] e 8 [µg/m³] per i LCS e 2 [µg/m³] e 7 [µg/m³] per il riferimento, dove viene sottolineato come questa corrispondenza sia legata al fatto che la zona di misura disti 2 [km] dalla stazione di riferimento.

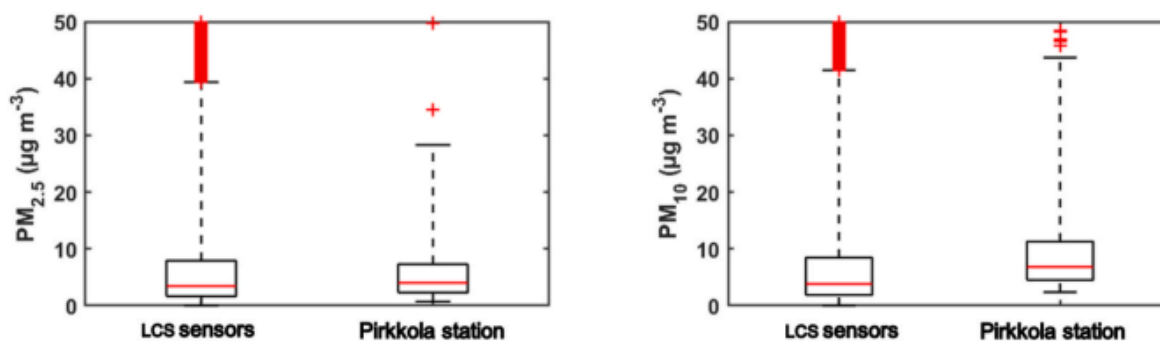


Figura 14: Boxplot PM_{2.5} e PM₁₀; confronto tra LCS e stazione di riferimento

Inoltre viene fatto notare come le misure di PM_{2.5} e PM₁₀, indoor e outdoor, seguano trend simili – come mostrato in Figura [15].

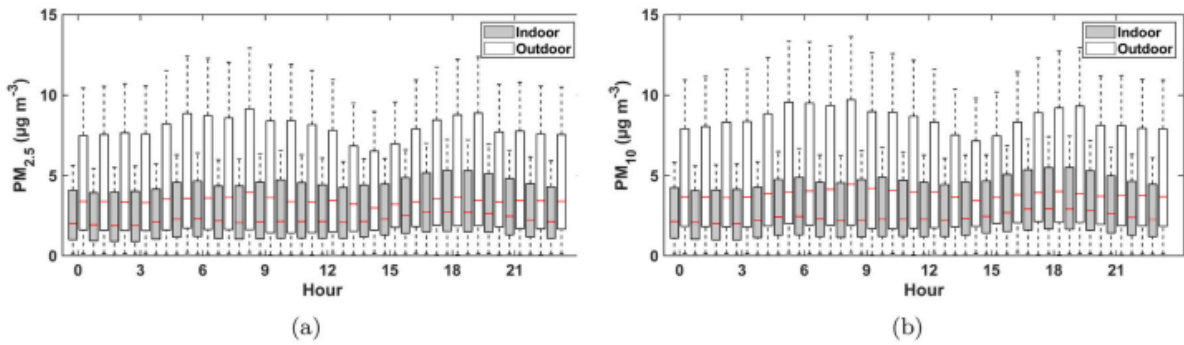


Figura 15: Boxplot dei trend di PM2.5 e PM10 indoor e outdoor

Nello studio si sottolinea come questa corrispondenza tra PM2.5 e PM10 permetta di identificare i periodi della giornata nei quali l'inquinamento dell'aria è maggiore, consentendo di comprendere quali possano essere le fonti delle emissioni. Si specifica infatti come i picchi delle concentrazioni siano tra le 7 e le 9 di mattina e tra le 16:30 e le 18 di sera, orari in cui è presente un maggiore traffico per via dei veicoli che vanno o ritornano dal lavoro. Inoltre, si nota come le misure indoor siano in media più basse di quelle outdoor, ma l'andamento è analogo. Ciò è spiegato dalla presenza di infiltrazioni dovute dalla scarsa coibentazione e dall'apertura delle finestre, dimostrando – come spiegato in studi precedenti – il 30-75% dell'inquinamento indoor derivi da quello outdoor.

Le *heatmap* presenti in Figura [16] dimostrano inoltre come l'utilizzo di LCS possa effettivamente offrire una copertura spaziale capillare, per permettere una comprensione dell'inquinamento e dell'esposizione personale più dettagliata.

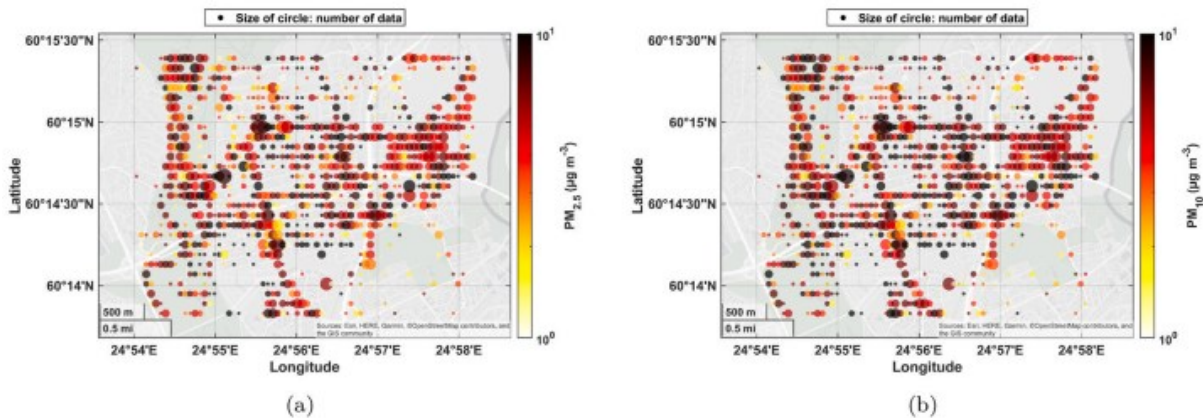


Figura 16: Heatmap PM2.5 e PM10

La tabella [6] mostra come l'esposizione personale sia fortemente dipendente dalle attività e abitudini dell'utente e dalla vicinanza ad aree verdi o trafficate – in quanto è possibile notare come i valori corrispondenti a ogni sensore riportato siano differenti tra loro.

Deposited Dose (μg)	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	S ₇	S ₈
Indoor (In)								
Male	0.770518	0.660556	0.942487	1.155383	1.976325	0.726582	1.619197	0.360656
Female	0.997141	0.854837	1.219689	1.495201	2.557597	0.940283	2.095431	0.466732
Outdoor (Out)								
Male	2.958538	2.270792	0.947633	2.673501	1.610770	1.171055	3.227640	2.353517
Female	3.828697	2.938672	1.226349	3.459825	2.084526	1.515483	4.176946	3.045728

Tabella 6: Esposizione personale al PM2.5

Capitolo 4: Materiali e metodi

In questo capitolo viene descritta e presentata la metodologia utilizzata nel progetto SOCIO-BEE per la misura della qualità dell'aria e l'analisi svolta sui dati raccolti.

Si illustra prima di tutto la piattaforma (Sezione 4.1.1.1) e l'applicazione (Sezione 4.1.1.2) del progetto SOCIO-BEE, dalla quale sono stati estrapolati i dati raccolti nelle campagne di misura. Successivamente vengono presentate le campagne di misura svolte dai cittadini – distinguendo tra misure da terra e misure con drone – delle tre città pilota, definendo prima il set-up sperimentale e poi i dati disponibili. Infine si mostra l'analisi dati svolta per ciascuna delle tre città pilota. Si guardi Figura [17] per avere una panoramica del procedimento dello studio.

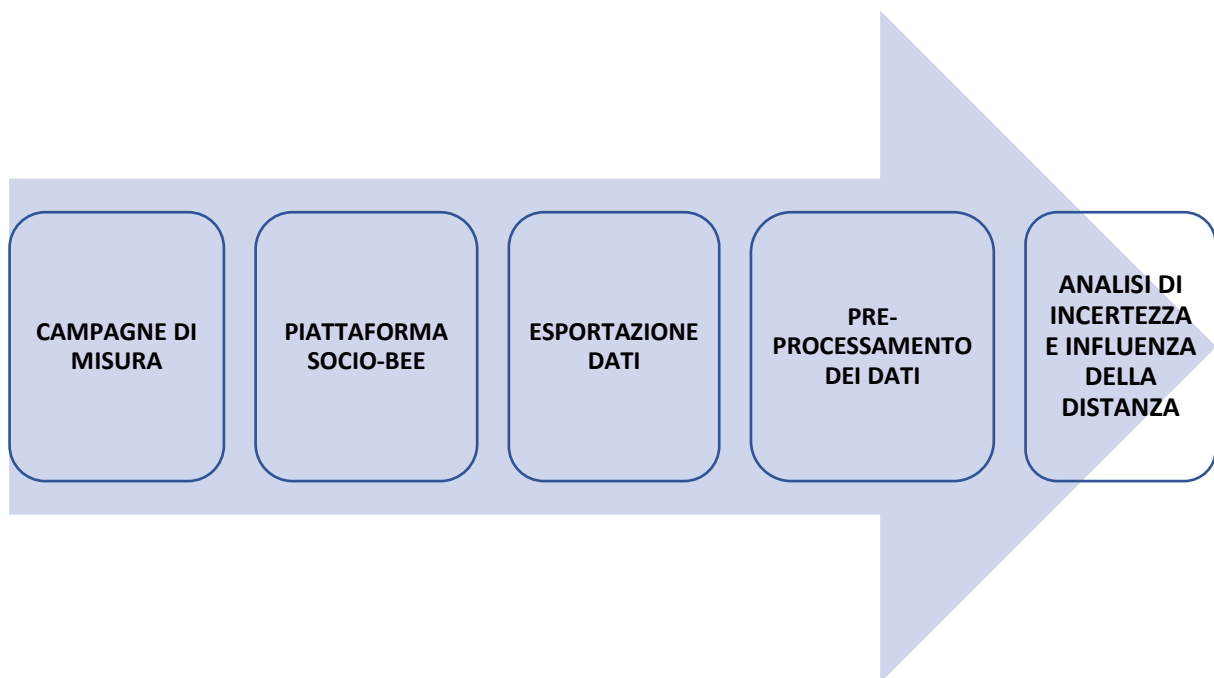


Figura 17: Fasi del procedimento dello studio.

4.1 Campagne di misura

Le tre città pilota – Ancona, Maroussi e Saragozza – sono state selezionate per condurre campagne di monitoraggio della qualità dell'aria. In queste città, sono stati distribuiti sensori a basso costo ai cittadini, che hanno contribuito effettuando misurazioni in specifiche aree urbane. La progettazione delle campagne di misura nelle città pilota ha seguito un processo ben definito: inizialmente, sono stati installati sensori fissi nelle aree designate; successivamente, sono stati testati i sensori a basso costo forniti da Bettair Cities. Dopo questa fase, è avvenuto il reclutamento dei volontari, i Citizen Scientists. Per prepararli, sono stati organizzati workshop informativi sul progetto e sui compiti da svolgere. Una volta conclusa la

formazione, si sono svolte le campagne di misura programmate. I dati raccolti sono stati poi analizzati e ne sono state tratte le relative conclusioni, che sono state infine utilizzate per attività di divulgazione sul tema ambientale. Il processo descritto è riassunto nella Figura [18].

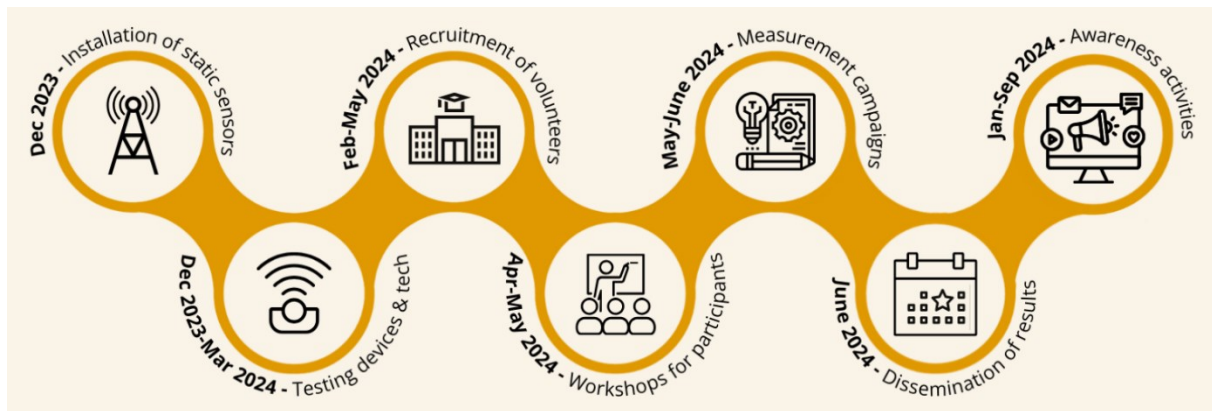


Figura 18: Fasi del procedimento del progetto nelle città pilota

4.1.1 Set-up sperimentale

Per meglio comprendere lo svolgimento delle campagne, si riprende la *Metafora dell'Alveare* enunciata nel Capitolo 1, andando a fornire informazioni più dettagliate sui compiti dei partecipanti – estrapolate dalla Derivable 5.7 del progetto SOCIO-BEE [23].

- **Apicoltore**, responsabile della creazione dell'alveare e dell'aggiunta di membri ad esso. È lui a creare le campagne seguendo le direttive del tool-kit di SOCIO-BEE.
- **Ape regina**, responsabile di configurare le campagne (scegliendo e.g. la zona, la metodologia di misura, la frequenza e il periodo) e di monitorare il loro corretto svolgimento. Inoltre prende parte alla pubblicazione ad interim dei risultati per aggiornare l'alveare e le parti interessate. Infine, è responsabile della chiusura delle campagne e della divulgazione dei risultati.
- **Api operaie**, responsabili di svolgere le misure durante le campagne, utilizzando i sensori indossabili – connessi ad un'applicazione sul cellulare – che gli sono stati precedentemente consegnati, seguendo le direttive dell'ape regina sulla zona da monitorare. Possono controllare regolarmente l'avanzamento della campagna e visualizzare *heatmap* realizzate con i dati raccolti. Una volta terminata la campagna, forniscono *feedback* sulla loro esperienza attraverso l'applicazione.
- **Api droni**, responsabili di raccogliere le informazioni sulle campagne e – formattandole in maniera diretta e comprensibile – di pubblicare risultati e conclusioni per permettere una migliore comprensione e maggiore consapevolezza sull'inquinamento dell'aria e le possibili azioni mitiganti.

Le differenze tra le tre città pilota sono da ricercare nel target di Citizen Scientists (anziani, giovani, pendolari) e nelle modalità delle campagne svolte. Perciò – dopo aver presentato la strumentazione utilizzata nelle campagne – si vanno a presentare i tre pilot, descrivendone i metodi.

Le modalità con cui sono state svolte le campagne sono state molteplici. Esse si possono suddividere in campagne di misura da terra – nelle quali i LCS sono stati distribuiti a cittadini; campagne di misura con drone – nelle quali i LCS sono stati montati su droni e sono state svolte misure a diverse altitudini; campagne di misura con autobus – nelle quali i LCS sono stati montati su autobus che percorrevano determinate tratte.

4.1.1.1 La piattaforma SOCIO-BEE

La piattaforma AcadeMe [25] utilizzata da SOCIO-BEE ha avuto una duplice funzione all'interno del progetto: la prima, raccogliere i dati raccolti durante le campagne, svolgere una breve analisi statistica e rappresentarli attraverso *heatmap* e grafici; la seconda, permettere agli utenti di interfacciarsi con le campagne, mostrando le *feature* appena elencate. All'interno della piattaforma è possibile visualizzare:

- Per le campagne attive, ovvero quelle in corso (Figura [19]):
 - Numero totale di campagne
 - Numero totale di utenti
 - Numero medio di utenti per campagna
 - Distanza totale coperta in [km]
 - Distanza media percorsa dagli utenti
 - La copertura media percentuale di ogni campagna
 - Il numero medio di sessioni di raccolta dati per campagna
 - La mappa contenente tutte le campagne
- Per le campagne completate, in aggiunta alle informazioni appena elencate:
 - Grafici con informazioni su: campagna, partecipanti e misure
 - Possibilità di filtrare tali informazioni per data

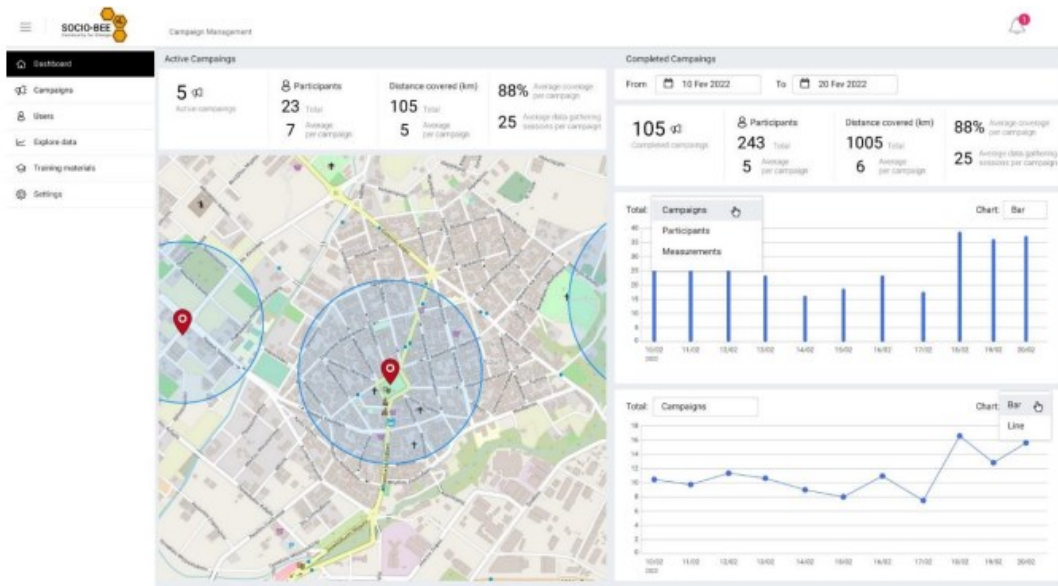


Figura 19: Piattaforma SOCIO-BEE; campagne attive.

Inoltre, viene mostrata una panoramica di tutte le campagne (Figura [20]): in essa è possibile selezionare la campagna desiderata per avere informazioni a riguardo, per quelle completate e in svolgimento, o per prenderne parte – per quelle in programma (Figura [21]).

Thumbnail	Title	Location	Date created	Status
	Thermi	Thermi square	18/02/2022	COMPLETED
	Marousi	Marousi	18/02/2022	ACTIVE
	Galerius	Ecoener	18/02/2022	PUBLISHED
	Thermi	Thermi square	18/02/2022	SAVED
	Marousi	Marousi	18/02/2022	PROPOSED
	Galerius	Ecoener	18/02/2022	DECLINED

Figura 20: Piattaforma SOCIO-BEE; panoramica campagne

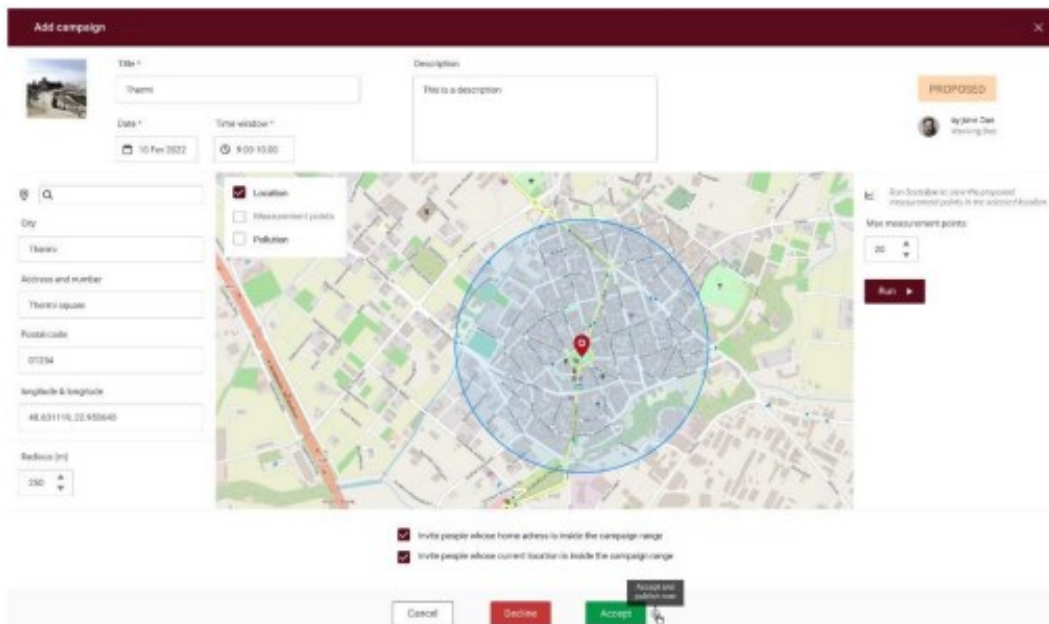


Figura 21: Piattaforma SOCIO-BEE; adesione a campagna

Nella piattaforma è possibile infine accedere a materiali formativi, come mostrato in Figura [22].

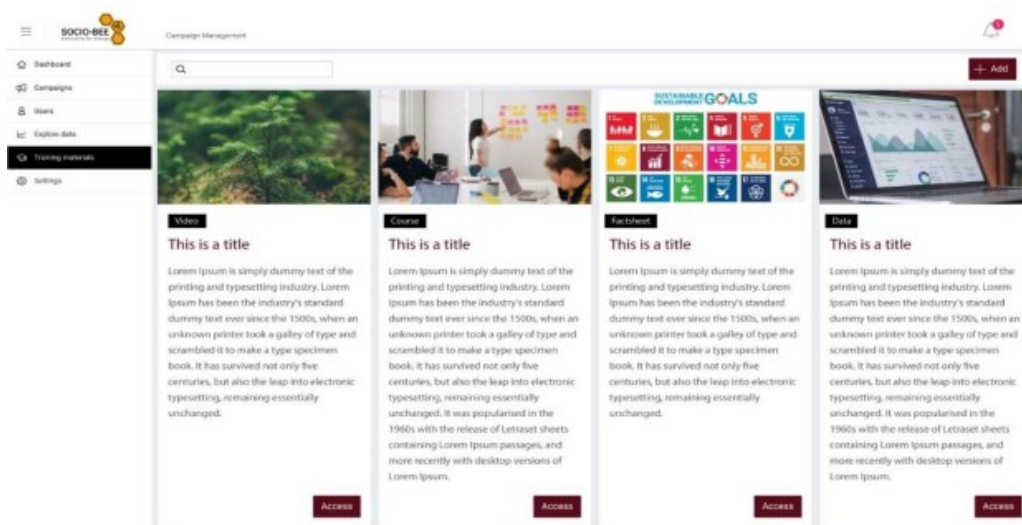


Figura 22: Piattaforma SOCIO-BEE; materiale formativo

4.1.1.2 L'applicazione SOCIO-BEE

L'applicazione di AcadeMe [25] utilizzata da SOCIO-BEE è stata, assieme alla piattaforma, lo strumento che ha permesso ai Citizen Scientists di interfacciarsi con il progetto. Grazie a essa infatti è stato possibile avere accesso ad informazioni riguardanti il numero di campagne a cui

si è preso parte, il numero di misure svolte, la *personal exposure* e l'impronta degli inquinanti – come si può vedere in Figura [23].

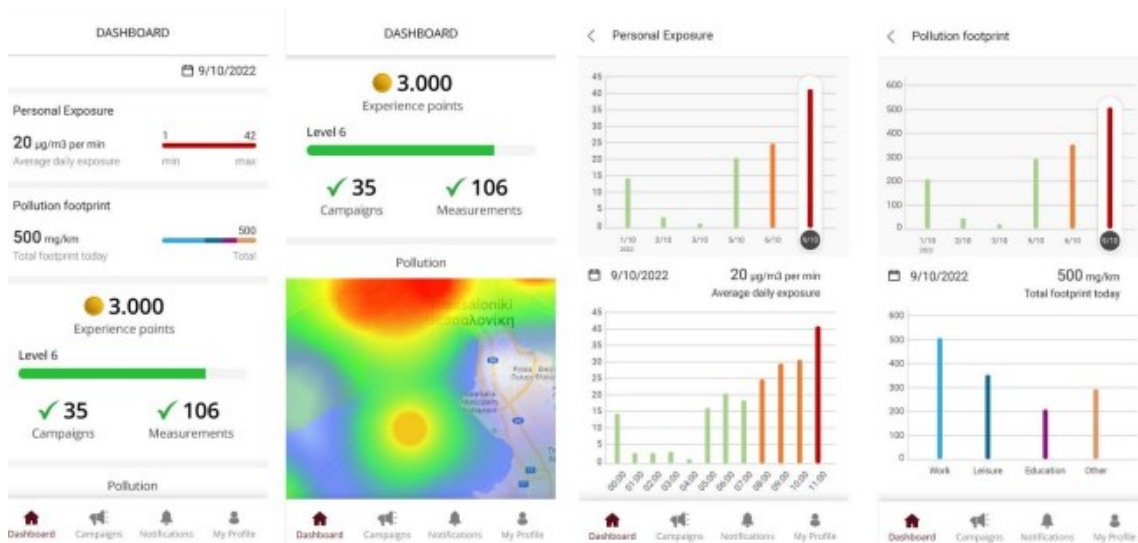


Figura 23: Applicazione SOCIO-BEE; dashboard e informazioni riguardo gli inquinanti

È inoltre possibile prendere visione delle campagne attive, completate e proposte, con la possibilità di ottenere informazioni riguardo l'area di interesse, la data e ora e la descrizione o il processo di raccolta dei dati durante una misura. Infine, l'applicazione dà la possibilità di visualizzare le campagne proposte e di scegliere se partecipare o meno, come mostrato in Figura [24].

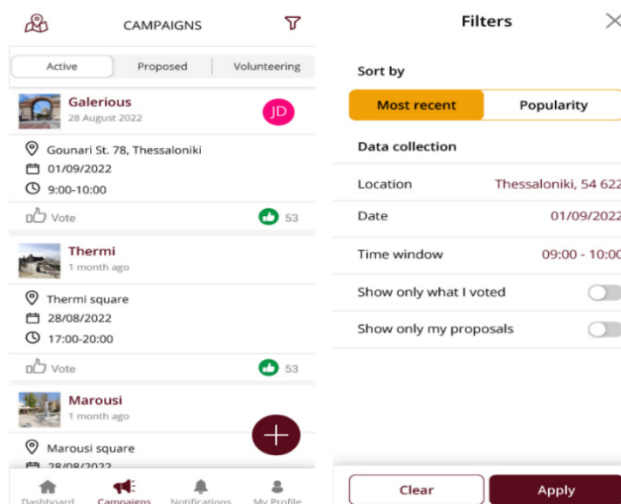


Figura 24: Applicazione SOCIO-BEE; campagne proposte

4.1.1.3 Wearable Sensor Node (WSN)

Il sensore utilizzato nel progetto di CS di SOCIO-BEE è stato fornito dall'azienda BETTAIR [24]. Il WSN è composto da un *Main Controller Module* (MCM) con incluso il modulo di connessione, un Sistema di Conversione analogico-digitale (SCa-d), batteria e quattro sensori (due integrati:

temperatura e umidità relativa, PM; due connessi attraverso i rispettivi moduli: O₃, NO₂). Nella Tabella [7] è possibile visualizzare il MCM e il SCA-d selezionati per il progetto, evidenziati in grassetto.

Component	Applicable TS	Compliant options *	Evaluation (1 is best)			
			Cost	Performance	Power	Area
Main Controller Module (including communications module)	TS20, TS24, TS25, TS26, TS29, TS31, TS32, TS33, TS34, TS36	ESPRESSIF ES32-C3 based	1	2	3	2
		ARM Cortex M0 based (ISP1302)	2	-1	1	1
		Xtensa LX6 dual core @ 240MHz based (Nina)	3	1	4	2
		ARM Cortex M4 based (ISP4520)	4	3	2	2
System Analog – to - digital converter ADC	TS9, TS27, TS28, TS29, TS30	ADS1115	1	3	1	1
		ADS1219**	2	2	2	2
		MCP3564	3	2	2	2
		ADS131M	4	2	2	3
		AD7173	5	1	3	4

-1: indicates non-compliance to Technical Spec *: not exhaustive **: Unavailable at the time of design

Tabella 7: Componenti del sensore proposti

Il MCM ESP32-C3-MINI-U1, rappresentato in Figura [25] all'interno della *main board*, include un'antenna e un *communication stack* per il Bluetooth 5 a 2.4GHz Wi-Fi (802.11 b/g/n).

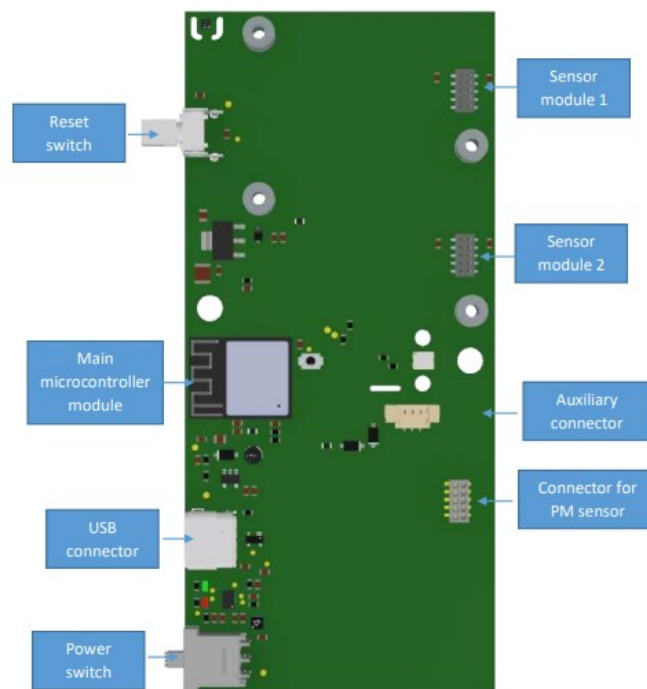


Figura 25: Main board WSN

In Figura [26] e Figura [27] sono rappresentati rispettivamente i grafici della potenza di trasmissione in funzione della distanza, del Bluetooth LE e del Wi-Fi (802.11 Standard).

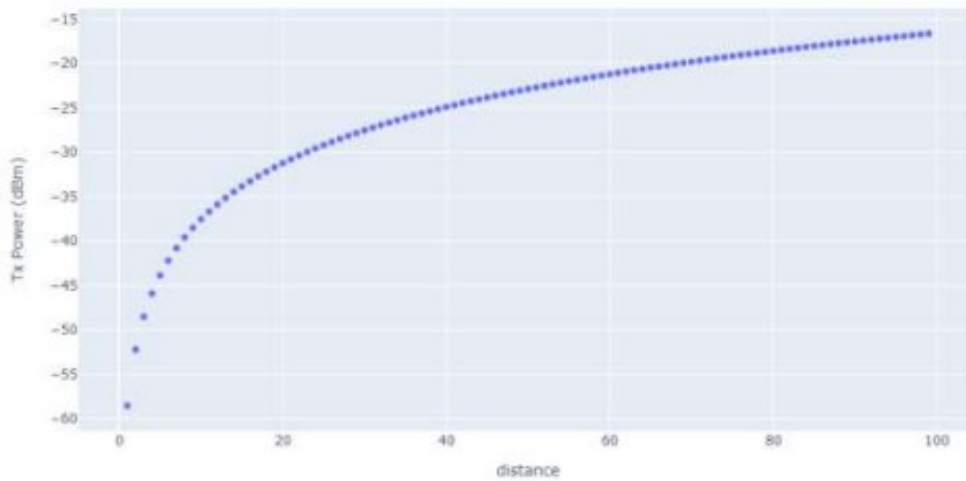


Figura 26: Grafico della potenza di trasmissione in funzione della distanza; Bluetooth

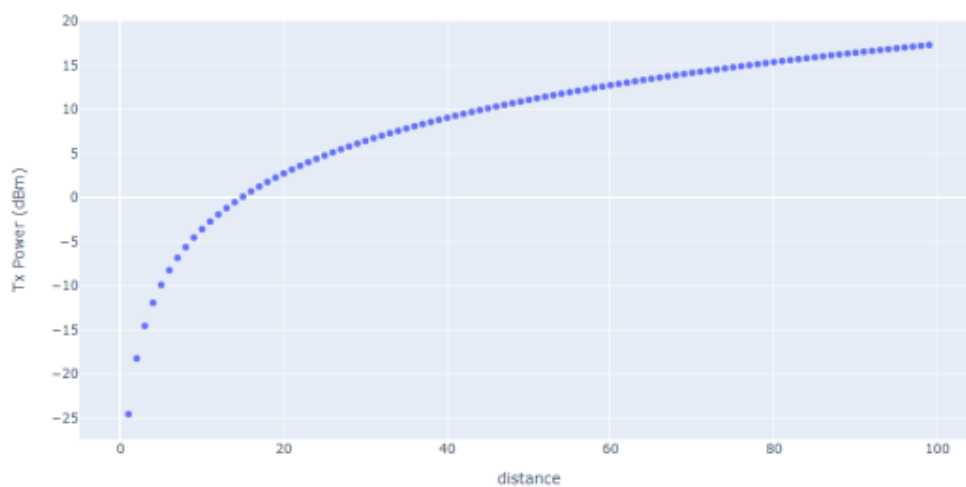


Figura 27: Grafico della potenza di trasmissione in funzione della distanza; Wi-Fi

La batteria selezionata è VARTA EasyPack, per la facilità di connessione con il WSN, la sicurezza, la capacità e il peso. Sono mostrate le specifiche nella Tabella [8].

ELECTRICAL SPECIFICATION	
Nominal voltage	3.7 V
Rated capacity at (0.5 C / 0.2 C), 23 °C ± 5 °C	630 mAh min., 660 mAh nominal
Watt-hour rating	2.5 Wh
Charging method	Constant current + constant voltage
Max. charge voltage	4.2 V (± 0.05 V)
Max. continuous charge current	630 mA
Rec. charge cut off	By current 6.3 mA or timer 3.5 h
Max. continuous discharge current	1,260 mA (limited by cell)
Rec. discharge cut off	3 V
Internal impedance	Approx. 115 mΩ
Exp. cycle life at (1.0 C / 1.0 C), 23 °C ± 5 °C	≥ 500 cycles ≥ 70 %

Tabella 8: Specifiche della batteria VARTA EasyPack

Il sensore integrato per la temperatura (T) e l'umidità relativa (RH) è il SHT40, con un'incertezza di base e range rispettivamente di 1,8% e 0/100% per RH e di 0,2°C e -40/+125°C per T. Il sensore integrato del particolato fine (PM2.5) è il PM2012, le quali specifiche possono essere visualizzate nella Tabella [9].

Laser Particle Sensor Specification	
Operating principle	Laser scattering
Measured particle range	0.3µm~10µm
Measurement range	0~5,000µg/m³
Resolution	1µg/m³
Working condition	-10°C ~ 60°C, 0-95%RH (non-condensing)
Storage condition	-40°C ~ 80°C, 0-95%RH (non-condensing)
Accuracy	PM2.5/PM1.0: 0~100µg/m³ , ±10µg/m³ 101~500µg/m³ , ±10% reading
PM1.0	
PM2.5	PM10: 0~100µg/m³ , ±25µg/m³ 101~500µg/m³ , ±25% reading
PM10	(GRIMM, 25±2°C, 50 ±10%RH)
Response time	1sec
Time to first reading	≤ 8 seconds
Power supply	DC 5V±0.1V, ripple wave <50mV
Working current	≤100mA
Standby current	≤200µA
Dimensions	W38*H35*D12 mm
Digital output 1 (default)	UART/IIC: 3.3V/5V(Level)
MTTF	128,000 hrs (continuous turn on)

Tabella 9: Specifiche del sensore PM2.5

Come possiamo notare dalla prima riga della Tabella [9], è un sensore che sfrutta il metodo ottico, che segue quindi il principio di funzionamento descritto nella Sezione 2.3.2.

I sensori di NO₂ e O₃ sono entrambi basati sul metodo elettrochimico (si veda Sezione 2.3.3 per il principio di funzionamento). Le specifiche sono visualizzabili in Tabella [10].

Analyte	LOD	Full Scale	Calibrated initial accuracy	Response time	Weight
NO ₂	<5ppb	1 ppm	5ppb	80s	6 gr
O ₃	<5ppb	1 ppm	5ppb	80s	6 gr

Tabella 10: Specifiche dei sensori NO₂ e O₃

Vengono ora presentati le dimensioni e un render con tutti i moduli inseriti del WSN in Figura [28].



Figura 28: Dimensioni e render del WSN

4.1.2 Ancona

Nella città pilota di Ancona, il target del progetto di CS sono stati gli anziani, con l'obiettivo di incoraggiarli a svolgere attività fisica all'aria aperta, in un ambiente non inquinato e non affollato, promuovendo uno stile di vita sano.

Sono state svolte in totale 6 campagne di misura da terra, in un lasso di tempo dalla durata di un mese, dal 27 maggio 2024 al 27 giugno 2024. Gli utenti che hanno partecipato a tali campagne sono stati 10, che hanno realizzato un totale di 225 misure. Nella Tabella [11] è possibile visualizzare un'analisi riassuntiva di tali campagne.

Analisi riassuntiva delle campagne	
NUMERO UTENTI	10
NUMERO CAMPAGNE	6
DURATA TOTALE CAMPAGNE	131 giorni
DURATA MEDIA CAMPAGNE	21 giorni e 22 ore
MEDIA NUMERO DI MISURE PER UTENTE	22.5
MEDIA NUMERO DI CAMPAGNE PER UTENTE	3.1
MEDIA NUMERO DI MISURE PER CAMPAGNA	37.5
MEDIA NUMERO UTENTI PER CAMPAGNA	5.2

Tabella 11: Analisi riassuntiva delle campagne di Ancona

In Figura [29] è possibile vedere dove sono state svolte le misure durante le campagne da terra e dei 3 sensori fissi posti al Comune di Ancona.



Figura 29: Ubicazione delle misure delle campagne da terra e dei sensori fissi

Le 3 campagne di misura con drone svolte in Ancona hanno raccolto un totale di 151 misure, visualizzabili in Figura [30]. In queste campagne sono stati utilizzati sensori SNIFFER, diversi da quelli utilizzati nelle campagne da terra, per via della migliore trasmissione a distanza delle misure. Le basi dei droni per le 3 campagne sono segnate in Figura [30].

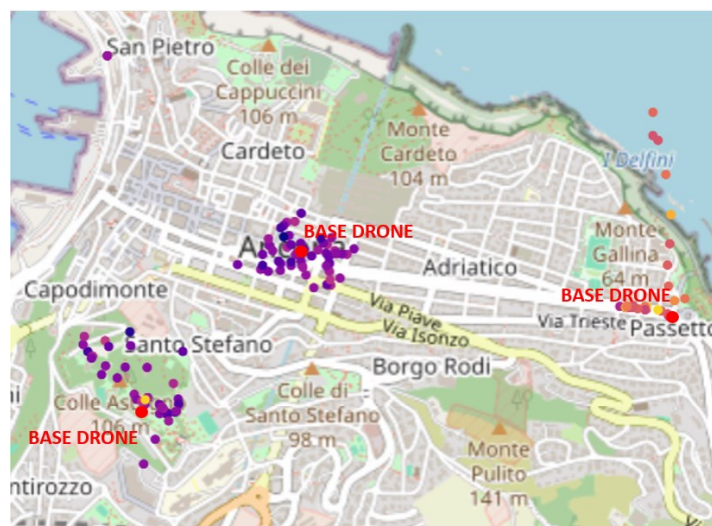


Figura 30: Ubicazione misure con drone

Per Ancona è stato possibile inoltre utilizzare i dati delle stazioni di riferimento di fondo, ottenuti dal sito di ARPAM (link: <https://aria.arpa.marche.it/Charts.aspx>), mostrate in Figura [31].



Figura 31: Ubicazione stazioni di riferimento

4.1.3 Maroussi

Nella città pilota di Maroussi, il target del progetto di CS sono stati i pendolari, con l'obiettivo di far comprendere l'importanza della qualità dell'aria e la responsabilità collettiva nel mantenere un ambiente sano.

Sono state svolte in totale 38 campagne di misura da terra, in un lasso di tempo dalla durata di poco più di un anno, dal 6 giugno 2023 al 28 giugno 2024. Gli utenti che hanno partecipato a tali campagne sono stati 61, che hanno realizzato un totale di 1206 misure. Nella Tabella [12] è possibile visualizzare un'analisi statistica riassuntiva di tali campagne.

Analisi riassuntiva delle campagne	
NUMERO UTENTI	61
NUMERO CAMPAGNE	38
DURATA TOTALE CAMPAGNE¹	57 giorni
DURATA MEDIA CAMPAGNE	36 ore
MEDIA NUMERO MISURE PER UTENTE	19.8
MEDIA NUMERO CAMPAGNE PER UTENTE	2.2
MEDIA NUMERO MISURE PER CAMPAGNA	31.7
MEDIA NUMERO UTENTI PER CAMPAGNA	3.5

Tabella 12: Analisi riassuntiva delle campagne di Maroussi

In Figura [32] è possibile vedere l'ubicazione delle misure svolte durante le campagne da terra e dei sensori fissi.

¹ somma delle durate di tutte le campagne

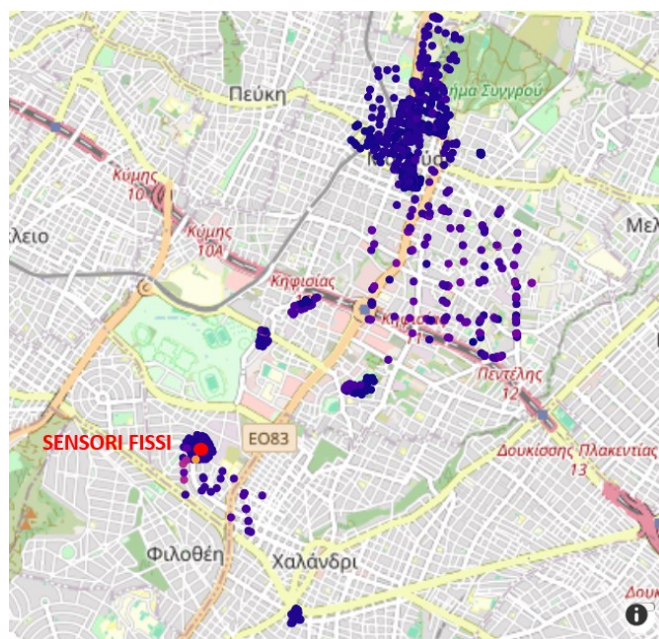


Figura 32: Ubicazione delle misure delle campagne da terra

4.1.4 Saragozza

Nella città pilota di Saragozza, il target del progetto di CS sono stati i giovani, con l'obiettivo di sensibilizzarli sulla tematica ambientale e far comprendere loro la possibile influenza che possono avere nel controllo e tutela della qualità dell'aria.

Sono state svolte in totale 13 campagne di misura da terra, in un lasso di tempo dalla durata di circa un mese, dal 20 giugno 2024 al 19 giugno 2024. Gli utenti che hanno partecipato a tali campagne sono stati 79, che hanno realizzato un totale di 405 misure. Nella Tabella [13] è possibile visualizzare un'analisi statistica riassuntiva di tali campagne.

Analisi riassuntiva delle campagne	
NUMERO UTENTI	79
NUMERO CAMPAGNE	13
DURATA TOTALE CAMPAGNE²	4 ore
DURATA MEDIA CAMPAGNE	18 minuti
MEDIA NUMERO MISURE PER UTENTE	5.1
MEDIA NUMERO CAMPAGNE PER UTENTE	1.5
MEDIA NUMERO MISURE PER CAMPAGNA	31.2
MEDIA NUMERO UTENTI PER CAMPAGNA	9.2

Tabella 13: Analisi riassuntiva delle campagne di Saragozza

In Figura [33] è possibile vedere l'ubicazione delle misure svolte durante le campagne da terra e dei sensori fissi.

² somma delle durate di tutte le campagne



Figura 33: Ubicazione delle misure delle campagne da terra

4.2 Analisi dati

Gli obiettivi di questo lavoro di tesi – come già menzionato nel Capitolo 1 – riguardano l'analisi della qualità dei dati raccolti tramite sensori a basso costo e l'analisi dell'influenza della distanza sulle misure. In quest'ottica, è stata condotta un'analisi sui dataset delle campagne di misura descritte in precedenza. L'elaborazione dei dati ha seguito due fasi principali: la prima ha riguardato il filtraggio e la validazione dei dati grezzi, mentre nella seconda sono stati applicati metodi matematici per identificare tendenze e anomalie. Per l'analisi dei dati è stato utilizzato il linguaggio di programmazione Python, con il supporto di Excel per l'archiviazione e l'organizzazione delle informazioni in formato tabellare. A causa della diversa disponibilità di dati – come discusso nella Sezione 4.1 – l'analisi è stata differente per ciascuna delle città pilota. La Tabella [14] fornisce una panoramica sulla disponibilità e comparabilità temporale dei dati, mentre la Tabella [15] riassume in modo generale quanto descritto nei capitoli successivi.

	ANCONA	MAROUSSI	SARAGOZZA
STAZIONI DI RIFERIMENTO			
SENSORI FISSI			
BUS			
CAMPAGNE			
DRONE			

	Dati disponibili
	Dati disponibili ma non comparabili
	Dati non disponibili
	Misurazioni non avvenute

Tabella 14: Panoramica sui dati delle tre città pilota

	ANCONA	MAROUSSI	SARAGOZZA
FILTRAGGIO			
ANALISI DI INCERTEZZA			
ANALISI GEOGRAFICA			

	Con riferimento
	Non svolta per incompatibilità dati
	Senza riferimento

Tabella 15: Panoramica sulle possibili analisi per le tre città pilota

Riassumendo quando rappresentato nella Tabella [15], si nota come i dataset delle tre città pilota presentano caratteristiche eterogenee, con Ancona che possiede il più completo, seguita da Maroussi e infine Saragozza. Ne consegue, come si nota nella Tabella [16], che per Ancona è stato possibile effettuare un'analisi completa, mentre per Maroussi e Saragozza è stata più limitata. Si sottolinea – come già descritto nella Sezione 4.1.1.1.1 – che i dati relativi alle campagne sono stati raccolti all'interno della piattaforma SOCIO-BEE e da lì scaricati. Per le stazioni di riferimento in Ancona – come descritto nella Sezione 4.1.1.2 – i dati sono stati scaricati dal sito ufficiale di ARPAM (Agenzia Regionale per la Protezione dell'Ambiente Marche) . Nelle sezioni successive vengono descritti le metodologie impiegate per l'analisi e le tecniche utilizzate per la modellizzazione e interpretazione dei dati di ciascuna città pilota.

4.2.1 Dataset e pre-processamento dei dati

Fase 1: Download dei dataset

I dataset, come già descritto, sono stati scaricati dalla piattaforma SOCIO-BEE, in formato CSV (Comma-Separated Values).

Fase 2: Caricamento dei dataset su Python

Su Python, grazie all'utilizzo della libreria *Pandas*, è possibile leggere file CSV. Dopo aver verificato il delimitatore e l'*encoding* dei dataset, essi sono stati caricati creandone dei *DataFrame* (DF). I DF delle campagne contano 11 colonne, 8 delle quali sono state utilizzate nel nostro studio: ID dell'utente, ID della campagna, latitudine, longitudine, *timestamp*, PM2.5, NO₂ e O₃. Per i DF dei sensori fissi sono invece state utilizzate 3 delle 7 colonne disponibili: ID del sensore, *timestamp* e PM2.5. Il DF delle campagne con drone conta 20 colonne, 5 quelle utilizzate: PM2.5, altitudine rispetto alla base, *timestamp*, latitudine e longitudine. Nelle Figure [34], [35] e [36] è possibile visualizzare un esempio di *dataframe* rispettivamente delle campagne da terra, dei sensori fissi e delle campagne con drone.

Index	campaign_id	ommendation	user_id	no2	o3	pm_2p5	r_humidity	temperature	timestamp	lat	lon
0	234	1396	238	84.4615	0	5.15385	43.4846	31.1785	1716808575000	43.6149	13.5306
1	234	1406	238	94.0769	0	6.53846	40.4623	32.2254	1716809243000	43.6154	13.5295
2	234	1415	10	15.8462	0	5.30769	41.7438	30.0692	1716823783000	43.6153	13.5307
3	234	1419	10	33.1538	0	4	43.4315	29.6	1716824222000	43.6159	13.5309
4	234	1431	10	18.1538	0	3.15385	42.2938	31.8123	1716826622000	43.6141	13.5332
5	234	1438	10	18.8462	0	3.07692	43.6192	30.5354	1716827061000	43.6146	13.5339
6	234	1505	10	33.6154	0	4.92308	41.7208	31.9338	1716883716000	43.6156	13.5335
7	234	1587	10	23.3077	0	6.46154	40.2838	33.4346	1716974158000	43.6153	13.5295
8	234	1591	10	34.6923	0	4.30769	44.9131	30.4462	1716974321000	43.6155	13.5298
9	234	1595	10	31.8462	0	6.46154	40.6169	32.7985	1716974497000	43.6153	13.5302
10	234	1596	10	41.5385	0	6.46154	41.3546	32.8423	1716974673000	43.6152	13.5307

Figura 34: Dataframe delle campagne da terra

Index	time	CARTRIDGE	wsnProcessed.mean_NO2	wsnProcessed.mean_O3	wsnProcessed.mean_PM_2P5	wsnProcessed.mean_R_HUMIDITY	wsnProcessed.mean_TEMPERATURE
0	2024-05-17T14:10:00.000Z	205FDD	22.1775	37.2895	1.2849	26.8836	33.8142
1	2024-05-17T14:10:00.000Z	218E05	12.4199	35.9891	1.78733	27.542	33.4903
2	2024-05-17T14:10:00.000Z	21A741	31.3885	16.5287	1.14177	27.0394	33.6283
3	2024-05-17T14:15:00.000Z	205FDD	22.052	36.9246	1.38154	26.8805	33.8568
4	2024-05-17T14:15:00.000Z	218E05	13.2269	34.3692	1.86416	27.5613	33.5188
5	2024-05-17T14:15:00.000Z	21A741	31.4318	16.2624	1.83828	26.9985	33.6788
6	2024-05-17T14:20:00.000Z	205FDD	21.5173	37.2125	1.61001	26.84	33.8795
7	2024-05-17T14:20:00.000Z	218E05	12.7803	34.7018	1.54461	27.5277	33.5338
8	2024-05-17T14:20:00.000Z	21A741	30.9055	16.3599	1.48716	26.9498	33.7148
9	2024-05-17T14:25:00.000Z	205FDD	21.219	37.1509	1.94656	26.8475	33.8918
10	2024-05-17T14:25:00.000Z	218E05	12.774	34.1038	1.68396	27.5283	33.5548

Figura 35: Dataframe dei sensori fissi

Index	Time Stamp	Abs.Alt m	Longitude	Latitude	Temperature °C	Humidity %	Pressure Pa	VOCs ppm	SO2 µg/m³	CO mg/m³	NO2 µg/m³	O3+NO2 µg/m³	PM1.0 µg/m³	PM2.5 µg/m³	PM10 µg/m³	CO2 mg/m³	Serial N°	named: 1	id_North n	wind_East m/s
0	2024-06-18 16:35:00	0	13.51...	43.6...	25.9125	48.6275	100680	0.	5.80396	79.3455	176.747	3.23077	6.15385	10.1538	1021...	853b6...	nan	nan	nan	
1	2024-06-18 16:36:00	0	13.51...	43.6...	26.1895	47.5392	100678	0.	5.68484	74.8387	170.686	3.21667	6.25	9.41667	1020...	853b6...	nan	nan	nan	
2	2024-06-18 16:37:00	0	13.51...	43.6...	26.6471	46.1601	100677	0.	5.50363	64.1822	158.832	2.86667	5.93333	9.48333	1020...	853b6...	nan	nan	nan	
3	2024-06-18 16:38:00	0	13.51...	43.6...	27.1424	44.6962	100677	0.	5.31362	59.446	147.154	2.90164	5.45902	9.42623	1019...	853b6...	nan	nan	nan	
4	2024-06-18 16:39:00	23.04...	13.51...	43.6...	25.317	46.1699	100420	0.	5.15861	38.6613	125.998	2.5	4.96667	8	1020...	853b6...	nan	nan	nan	
5	2024-06-18 16:40:00	34.85...	13.51...	43.6...	23.0915	49.1144	100244	0.	4.76659	7.48309	110.788	3.03333	6.06667	9.4	1004...	853b6...	nan	nan	nan	
6	2024-06-18 16:41:00	377895	13.51...	43.6...	22.9118	51.8333	100635	0.	4.31118	48.7615	161.492	2.85	5.58333	9.15	1023...	853b6...	nan	nan	nan	
7	2024-06-18 16:42:00	142771	13.51...	43.6...	23.7549	50.3137	100634	0.	4.12828	56.4483	167.1	3.13333	6.2	11.0833	1023...	853b6...	nan	nan	nan	
8	2024-06-18 16:43:00	38.17...	13.51...	43.6...	21.8235	53.5752	100231	0.	4.02762	25.419	157.492	2.71667	5.18333	8.3	1023...	853b6...	nan	nan	nan	
9	2024-06-18 16:44:00	43.43...	13.51...	43.6...	20.781	55.1928	100157	0.	3.6652	22.1985	159.645	2.16667	4.36667	7.36667	1025...	853b6...	nan	nan	nan	
10	2024-06-18 16:45:00	41.29...	13.51...	43.6...	19.768	57.1765	100175	0.	3.38036	30.2771	168.958	4	7.28333	10.2833	1026...	853b6...	nan	nan	nan	

Figura 36: Dataframe delle campagne con drone

4.2.2 Filtraggio preliminare dei dataset e analisi statistica

Successivamente al caricamento e pre-processamento dei dati, è stato effettuato un filtraggio preliminare dei dataset, eseguito con le seguenti fasi.

Fase 1: Visualizzazione dei dati

Utilizzando la libreria *matplotlib.pyplot* di Python, sono stati stampati i plot di ciascun inquinante in funzione dell'indice temporale, con l'obiettivo di ottenere una panoramica del dataset, visualizzando graficamente il trend e le anomalie più notevoli.

Fase 2: Filtraggio preliminare dei dataset

Per filtrare preliminarmente i dataset è stata utilizzata una soglia data dalla deviazione standard (σ , STDV) delle misure del dataset stesso moltiplicata per un fattore di copertura pari a 3. Per calcolare la STDV è stata adoperata la funzione *.std()* di Python che impiega la seguente formula.

$$\sigma = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n e_i^2} = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2}$$

con x_i = misurazione i -esima e \bar{x} = media delle misurazioni.

Dopo aver eseguito il filtraggio preliminare, si è valutato come gli inquinanti O₃ e NO₂ avessero dei dataset di qualità inferiore rispetto al PM_{2.5}, quindi l'analisi successiva si è concentrata solamente sul particolato fine. Dopo aver completato il pre-processamento dei dati, è stata realizzata una breve analisi statistica – i quali risultati si possono leggere nelle Sezioni dalla 4.1.1.2 alla 4.1.1.4 – utilizzando funzioni di raggruppamento, somma, calcolo media e conteggio su Python ed Excel. Una volta eseguite le Fasi 1 e 2 elencate nella Sezione 4.2.1 e quelle elencate nella presente Sezione, l'analisi è stata condotta – come già detto più volte – in maniera differente per ogni città pilota. Nelle Sezioni successive si presenta l'analisi svolta per Ancona, Maroussi e Saragozza.

4.2.3 Ancona: analisi dati

Per la città pilota di Ancona, riprendendo la Tabella [16], l'analisi ha seguito diverse fasi. Dopo aver eseguito i passaggi elencati nelle Sezioni 4.2.1 e 4.2.2, sono stati caricati i file contenenti i dati delle stazioni di riferimento e i dati delle misure dei sensori fissi posti al Comune e sono stati anch'essi pre-processati. Dopo aver preso visione dei dataset e dei plot delle misure dei riferimenti, si è valutato che la stazione di riferimento posta a Cittadella fosse la migliore per effettuare l'analisi, in quanto quella posta alla Stazione Ferroviaria presentava buchi temporali per periodi non brevi. In seguito, sono state condotte due tipologie di analisi: la prima, di incertezza – con lo scopo di indagare sulla qualità dei dati dei sensori indossabili; la seconda, geografica – con lo scopo di valutare l'influenza della distanza sulle misure. Si presentano ora le metodologie seguite per svolgere le due analisi appena elencate.

4.2.3.1 Analisi di incertezza

Per eseguire l'analisi di incertezza si sono confrontati i dataset delle campagne e dei sensori fissi con il dataset della stazione di riferimento di Cittadella, per poter valutare la performance del sensore a basso costo sotto più punti di vista, eliminando le variabili dovute alla non perfetta esecuzione della misura da parte dell'utente – nel caso delle campagne – e alla stazionarietà del sensore – nel caso dei sensori fissi. Si presentano ora le fasi dell'analisi in questione, svolta in maniera identica per le campagne e per i sensori fissi. Chiamiamo RIF il riferimento (stazione di Cittadella) e DAT il dataset da confrontare con il suddetto riferimento (nel primo caso le campagne di misura, nel secondo i sensori fissi). Si aggiunge che nel caso delle misure con sensori fissi, sono stati selezionati solo 2 dei 3 sensori disponibili, in quanto dalla visione del dataset e del grafico si è evidenziato un problema tecnico che ne ha inficiato la performance (per due terzi del periodo di funzionamento i dati presentano valore pari a 0).

Fase 1: Calcolo della media oraria di DAT

Per permettere il confronto con le misure di RIF, si è reso necessario un ulteriore processamento di DAT. La frequenza delle misure di RIF è infatti oraria, mentre per DAT è differente (il dataset delle campagne non segue alcun criterio temporale – essendo le misure avvenute per mano di vari utenti in periodi differenti; quello dei sensori fissi conta una misura ogni 5 minuti). Perciò è stata calcolata la media oraria delle misure per i due suddetti dataset.

Fase 2: Calcolo dei residui

Per procedere con l'analisi si sono calcolati i residui (Figura [40]) della media oraria di DAT rispetto a RIF, sottraendo colonna a colonna le misure di PM2.5.

$$e_i = x_{DAT,i} - x_{RIF,i}$$

Fase 3: Filtraggio rispetto RIF

È stato eseguito un secondo filtraggio, in aggiunta a quello preliminare, utilizzando come soglia la deviazione standard dei residui di DAT rispetto a RIF, con fattore di copertura pari a 3. Per meglio visualizzare il corretto funzionamento del filtraggio e l'eventuale necessità di un terzo (così non è stato), sono stati stampati quattro grafici: il primo, raffigurante RIF e DAT pre-filtraggio; il secondo, con RIF e DAT post-filtraggio; il terzo, con DAT pre e post filtraggio; il quarto, uno *scatter* di DAT e RIF con evidenziati i dati eliminati.

Fase 4: Analisi di regressione lineare

Per svolgere tale analisi è stata utilizzata la funzione *LinearRegression* di Python, della libreria *sklearn.linear_model*. Questa funzione prende in ingresso le colonne di RIF e DAT di PM2.5 e sfruttando il metodo dei minimi quadrati, va a calcolare i coefficienti della retta di regressione che minimizzano la deviazione tra RIF e DAT, seguendo questi passaggi:

- Il modello prende in ingresso due variabili: x la colonna di RIF e y la colonna di DAT.
- Utilizzando i valori i -esimi delle colonne x e y , il modello viene addestrato a predire i valori di y , che chiameremo \hat{y} . Viene quindi creata l'equazione della retta di regressione:

$$\hat{y} = b + mx$$

- Il modello va ora a calcolare i coefficienti che minimizzano l'errore quadratico medio (MSE) medio tra y e \hat{y} .

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

con $n =$ numero di misurazioni delle colonne x e y

- Una volta calcolati i coefficienti definitivi, il modello è ora in grado di prevedere correttamente i valori di y in funzione di x .

Fase 5: Calcolo di coefficienti per completare l'analisi

In un unico grafico, vengono stampati i valori di x e y e la retta di regressione con i coefficienti prima calcolati. Successivamente vengono calcolati: coefficiente di Pearson, media (μ) e STDV – fattore di copertura 2 – di DAT e RIF, media e STDV – fattore di copertura 2 – dei residui di DAT rispetto a RIF. Il coefficiente di Pearson r è una misura del grado di correlazione tra DAT e RIF [9] e si calcola utilizzando una funzione della libreria *numpy* di Python, che è rappresentata dalla seguente equazione:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2} \cdot \sqrt{\sum(y_i - \bar{y})^2}}$$

4.2.3.2 Analisi geografica

Per eseguire l'analisi geografica, è stato confrontato il dataset delle campagne – che chiamiamo CAM – con quello della stazione di riferimento di Cittadella – che chiamiamo ancora RIF. Anche in questa analisi, sono state utilizzate le Fasi 1, 2 e 3 descritte nella Sezione 4.2.3.1, per pulire CAM dagli *outlier*. Inoltre, è stato filtrato temporalmente anche RIF, in quanto – essendo l'attenzione posta sul legame tra distanza e misura – l'influenza del periodo di misura non è stata considerata: sono stati considerati *outlier* tutti i valori di RIF nelle fasce orarie al di fuori di quelle delle misure di CAM. Si presentano ora le fasi dell'analisi in questione.

Fase 1: Visualizzazione delle heatmap

Per visualizzare le misure di CAM in funzione della loro posizione spaziale è stata utilizzata una *heatmap*, creata grazie alla libreria *plotly.express* e stampata come file *.html* grazie alla libreria

webbrowser. In tale *heatmap*, passando il cursore su ogni misura, è possibile prendere visione di: data e ora, valore di PM2.5, latitudine e longitudine.

Fase 2: Calcolo della distanza media tra misure

Per il calcolo della distanza media tra misure è stata usata una funzione della libreria *geopy.distance*: con essa è stata calcolata la distanza tra una misura e la successiva temporalmente e poi ne è stata fatta la media.

Fase 3: Creazione di *cluster* geografici attraverso machine learning non supervisionato

Dopo aver preso visione delle misure sulla *heatmap*, è stato valutato di creare 3 cluster utilizzando la funzione *Kmeans* della libreria *sklearn.cluster*. Tale funzione minimizza la Somma delle Distanze Quadrate (SSE).

$$SSE = \sum_{k=1}^K \sum_{i \in C_k} \|x_i - \mu_k\|^2$$

K = numero di cluster

C_k = set di punti assegnati al cluster k

x_i = punto appartenente a C_k

μ_k = centroide del cluster k

$\|x_i - \mu_k\|^2$ = distanza euclidea al quadrato

La minimizzazione è ottenuta sfruttando l'algoritmo seguente:

- **Assegnazione ai cluster:** Ogni punto x_i viene assegnato al cluster il cui centroide μ_k è il più vicino, calcolando la distanza euclidea.

$$d(x_i, \mu_k) = \sqrt{\sum_{j=1}^n (x_{ij} - \mu_{kj})^2}$$

- **Calcolo dei nuovi centroidi:** Una volta assegnati i punti ai cluster, si ricalcola il centroide di ogni cluster come la media dei punti assegnati a quel cluster.

$$\mu_k = \frac{1}{|C_k|} \sum_{i \in C_k} x_i$$

- **Ripetizione:** Si ripetono i passi 2 e 3 fino a che i centroidi non cambiano più.

Fase 4: Calcolo delle distanze e medie

Per completare l'analisi geometrica sono state calcolate: distanza tra i 3 centroidi dei *cluster*; distanza dei 3 centroidi dalle coordinate di RIF; media e STDV delle misure di RIF, di CAM e dei cluster.

4.2.3.3 Analisi dati delle campagne con drone

L'analisi dati delle campagne con drone – che si è limitata a due campagne, a Cittadella e a Piazza Cavour – è stata eseguita riprendendo per prima cosa tutte le Fasi descritte nella Sezione 4.2.1. ad ausilio delle misure con drone, sono state svolte negli stessi giorni delle campagne da terra aggiuntive da terra.

Tuttavia, prendendo visione dei plot in funzione dell'indice temporale, si è notata una non corrispondenza negli orari a cui sono state svolte le misure. Da ciò è derivata una limitatezza nell'analisi.

Si presentano le fasi dell'analisi in questione.

Fase 1: Visualizzazione delle *heatmap* delle campagne con drone

Per avere una panoramica delle misure e meglio comprendere la loro ubicazione, è stata creata una *heatmap* rappresentante tutte le misure di PM2.5 delle campagne con drone.

Fase 2: Creazione di grafici delle misure in funzione dell'indice temporale e dell'altezza sul livello del mare

Inizialmente si è calcolata l'altezza sul livello del mare delle misure, sommando alla colonna dell'altezza relativa alla base del drone l'altitudine della suddetta base.

- Base Cittadella: 98 m s.l.m
- Base Piazza Cavour: 66 m s.l.m

Una volta calcolata, è stato creato un grafico a tre variabili – utilizzando la libreria *seaborn* di Python – che mette in relazione il valore di PM2.5 all'indice temporale e all'altezza s.l.m.

Fase 3: Calcolo media e deviazione standard delle campagne con drone, delle campagne da terra ausiliarie e dei residui delle campagne rispetto al RIF

Per concludere l'analisi sulle campagne con drone, sono state calcolate le medie e STDV delle campagne, delle campagne con misure da terra ausiliarie e dei residui di esse rispetto a RIF.

4.2.4 Analisi dati: Maroussi

Per la città pilota di Maroussi, riprendendo la Tabella [16], l'analisi è stata limitata solo all'ambito geografico. Infatti, dopo aver eseguito i passaggi elencati nelle Sezioni 4.2.1 e 4.2.2, si è notato come non ci sia corrispondenza temporale tra le misure delle campagne e quelle dei sensori fissi. Per verificare questa ipotesi, è stato fatto un controllo incrociato utilizzando

la funzione `.intersection()` della libreria *pandas* tra gli indici temporali dei due dataset: da esso è emerso come le misure avvenute in periodi comparabili siano state solamente 11, un numero troppo esiguo per permettere un'ulteriore analisi. Quanto appena detto, in aggiunta alla non disponibilità dei dati delle stazioni di riferimento, non ha permesso di svolgere l'analisi di incertezza.

4.2.4.1 Analisi geografica

L'analisi geografica per la città di Maroussi ha seguito lo stesso approccio di quella svolta per Ancona, con l'unica differenza che essa si è limitata ai giorni con il maggior numero di misure, nei quali è stato possibile ottenere i dati di due stazioni di riferimento. Tale analisi è stata perciò eseguita mediante le seguenti fasi.

Fase 1: Selezione dei giorni con il maggior numero di misure

Per ottenere i giorni con più misure, inizialmente si è conteggiato – utilizzando una funzione della libreria *pandas* – il numero di dati raccolti giorno per giorno, ordinandoli in maniera decrescente. Successivamente, sono state selezionati i giorni con un numero di misure maggiore di 30, che è risultato in un totale di 10 giorni.

Fase 2: Ricerca dei dati delle stazioni di riferimento

Per i 10 giorni individuati, è stato possibile trovare le medie giornaliere di PM2.5 di due stazioni di riferimento sul sito <https://aqicn.org/map/athens/>

Fase 3: Ripetizione delle Fasi 1-4 dell'analisi geografica di Ancona

Per i 10 giorni individuati, è stata svolta la stessa analisi svolta per Ancona, dalla Fase 1 alla Fase 4, ottenendo quindi: distanza tra i 3 centroidi dei *cluster*; distanza dei 3 centroidi dalle coordinate di RIF; media e STDV delle misure di RIF, di CAM e dei cluster.

Fase 4: Ripetizione della Fase 5 dell'analisi geografica di Ancona per i due giorni con il miglior dataset

Presa visione delle grandezze calcolate nella Fase precedente, si è notato come solamente in due di questi giorni – il primo: 03/04/2024 e il quarto: 23/05/2024 – le medie dei 3 cluster differiscono di meno del 50% dalle medie dei riferimenti.

Perciò la Fase 5 dell'analisi geografica di Ancona è stata svolta solamente nei due giorni appena menzionati.

4.2.5 Analisi dati: Saragozza

Per la città pilota di Saragozza, riprendendo la Tabella [16], l'analisi è stata limitata solamente al pre-processamento, analisi statistica e filtraggio preliminare dei dati. Infatti, anche qui come per Maroussi, è stata verificata l'incompatibilità temporale dei dati delle campagne con quelli dei sensori fissi – sfruttando lo stesso metodo – ottenendo un totale di 16 misure comparabili temporalmente.

In aggiunta all'analisi preliminare, sono state calcolate media e STDV del dataset pre e post filtraggio, per prendere visione del miglioramento dei dati.

Capitolo 5: Risultati

Di seguito vengono presentati i risultati delle analisi descritte nel Capitolo 4, relativi alle tre città pilota.

5.1 Ancona

Per la città pilota di Ancona i risultati sono suddivisi in tre ulteriori sezioni, ciascuna inerente al tipo di analisi svolto per ottenerli.

5.1.1 Risultati dell'analisi di incertezza

Vengono qui presentati i risultati dell'analisi di incertezza, svolta con lo scopo di indagare la qualità dei dati raccolti, per valutare l'utilizzo dei sensori a basso costo descritti nella Sezione 4.1.1.1.3 all'interno del progetto di Citizen Science oggetto del nostro studio.

Per prima cosa, si tratta il confronto tra i sensori fissi e la stazione di riferimento a Cittadella. In Figura [37] sono presenti il grafico rappresentante gli andamenti delle misure dei sensori fissi a confronto con quelle dei riferimenti affiancato dal grafico pre e post filtraggio dei sensori fissi. È necessario sottolineare come il picco di quasi 80 $\mu\text{g}/\text{m}^3$ e le misure successive misurate dal 22/06/2024 dalle stazioni di riferimento sono da collegare ad un incendio avvenuto in un magazzino di smaltimento rifiuti avvenuto fuori dal centro città, dove erano posti i sensori fissi. Da essi si nota come il sensore segua precisamente il trend dei riferimenti, ma le misure siano in media sempre minori.

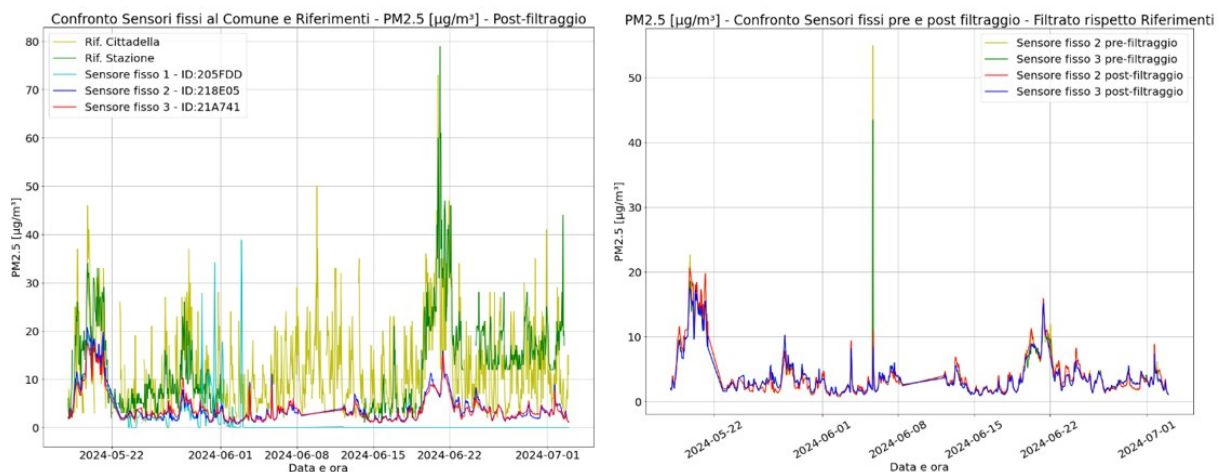


Figura 37: Grafici di confronto tra sensori fissi e stazioni di riferimento

In Figura [38] sono presentati i risultati dell'analisi di regressione tra sensori fissi e riferimento di Cittadella, dai quali è possibile trarre come conclusione l'effettivo miglioramento del modello di correlazione tra i dati delle campagne da terra rispetto quelli del riferimento di Cittadella. Essendo gli andamenti dei due sensori analizzati molto simili, per semplicità si presentano i risultati del confronto di uno solo dei due.

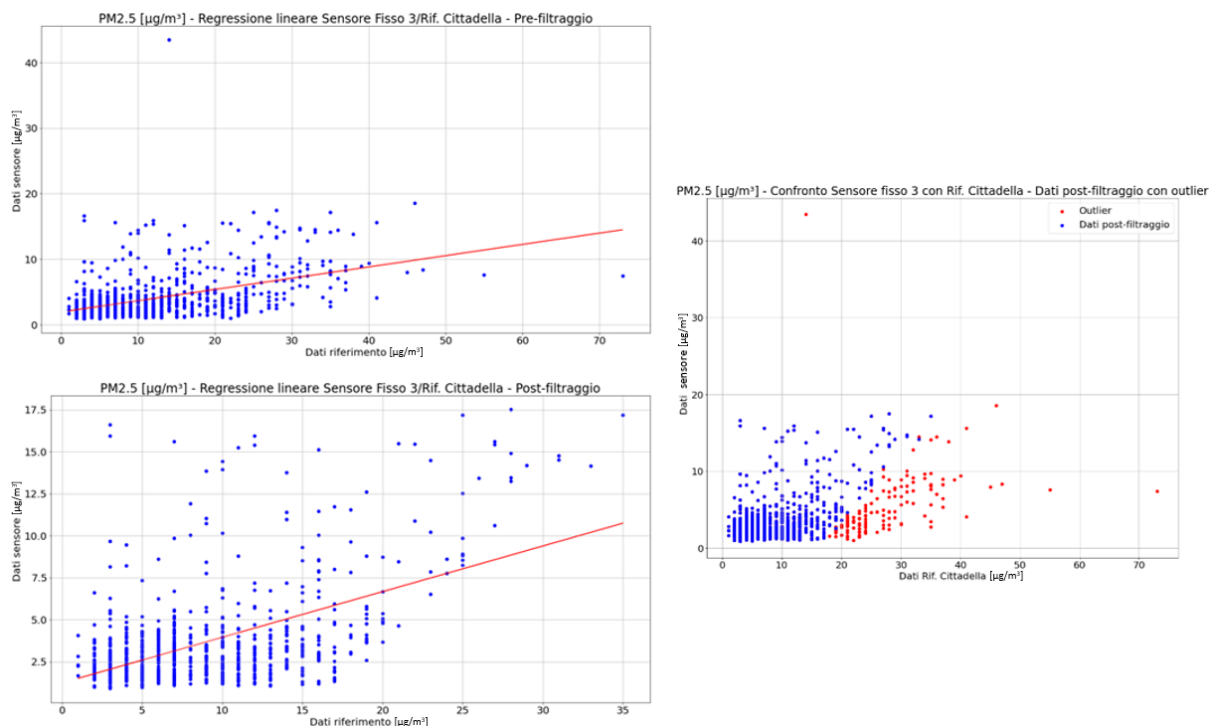


Figura 38: Grafici dell'analisi di regressione tra sensori fissi e riferimento di Cittadella; pre e post filtraggio

Nella Tabella [16] sono presentati i risultati complessivi del confronto tra le misure con i sensori fissi e il riferimento di Cittadella.

PM2.5 [$\mu\text{g}/\text{m}^3$] - CONFRONTO STAZIONE DI RIFERIMENTO E SENSORE FISSO	
Pendenza retta di regressione	0.27
Intercetta retta di regressione (Bias)	1.23
μ residui	5.33
σ residui	9.70
Coefficiente di Pearson	0.52
μ misure del riferimento	9.01
σ misure del riferimento	11.39
μ misure del sensore fisso	3.68
σ misure del sensore fisso	5.90

Tabella 16: Risultati dell'analisi di incertezza; confronto tra campagne di misura da terra e stazione di riferimento

Si nota una modesta correlazione lineare tra i dati dei sensori fissi e quelli del riferimento, in quanto il coefficiente di Pearson è pari a 0.52. Tuttavia questo dato non è da considerare come negativo, in quanto i fattori che possono aver contribuito a questo valore sono molteplici, primo fra tutti la distanza tra le due posizioni spaziali. Infatti, andando a considerare le deviazioni standard delle misure del sensore fisso e quella del riferimento, si evidenzia come la prima sia quasi la metà della seconda (5.90 contro 11.39 [$\mu\text{g}/\text{m}^3$]) – a dimostrare una minore dispersione delle misure. Degni di nota sono anche il valore della media dei residui (5.33 [$\mu\text{g}/\text{m}^3$]) e la pendenza della retta di regressione (0.27), che aiutano a quantificare quel discostamento costante tra i due andamenti delle misure. Infine, la deviazione standard dei

residui pari a $9.70 \mu\text{g}/\text{m}^3$ rivela un'elevata dispersione e quindi una possibile non accuratezza nella misura, ma sottolineando nuovamente l'estrema similarità di trend (si guardi Figura [37]), le cause di questo valore sono ancora una volta da ricercare – tra le altre cose – nella differente collocazione. Successivamente, si tratta il confronto tra le campagne di misura da terra e il riferimento di Cittadella. Nei grafici di Figura [39] è possibile visualizzare l'andamento delle misure dei due dataset a confronto e l'influenza del filtraggio. Si sottolinea che il picco di oltre $28 \mu\text{g}/\text{m}^3$ di PM_{2.5} del giorno 29/05/2024 sia stato misurato in una banchina del porto di Ancona in concomitanza con la partenza di un traghetto, perciò non è stato considerato come *outlier*.

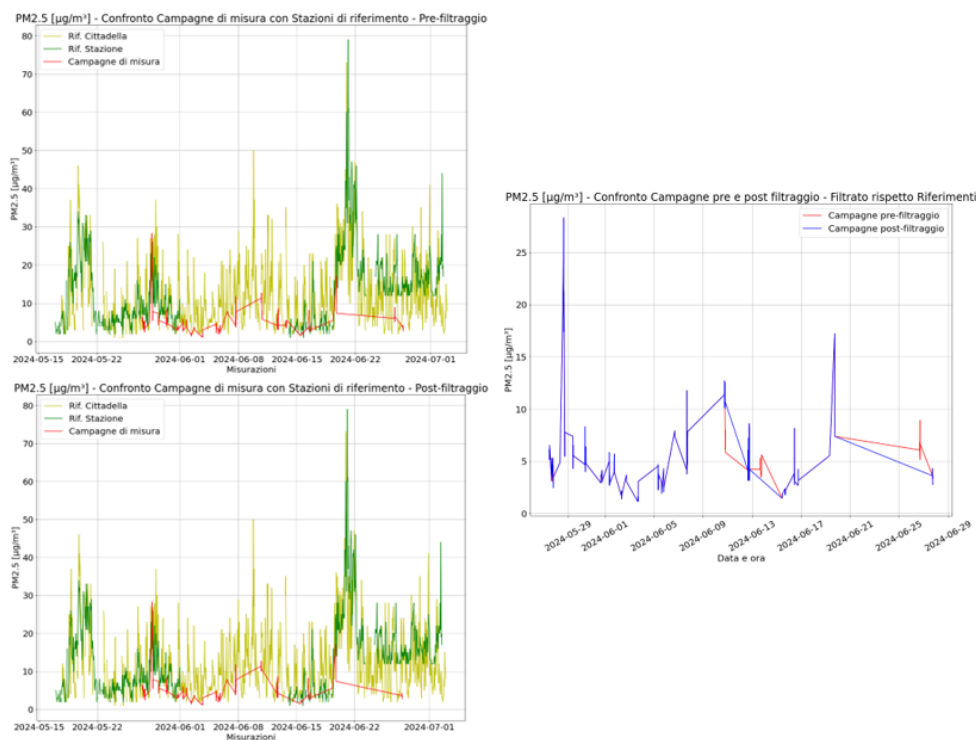


Figura 39: Grafici del confronto tra campagne di misura da terra e stazioni di riferimento; pre e post filtraggio

In Figura [40] sono invece presentati i grafici risultanti dall'analisi di regressione ai minimi quadrati e i dati eliminati mediante filtraggio.

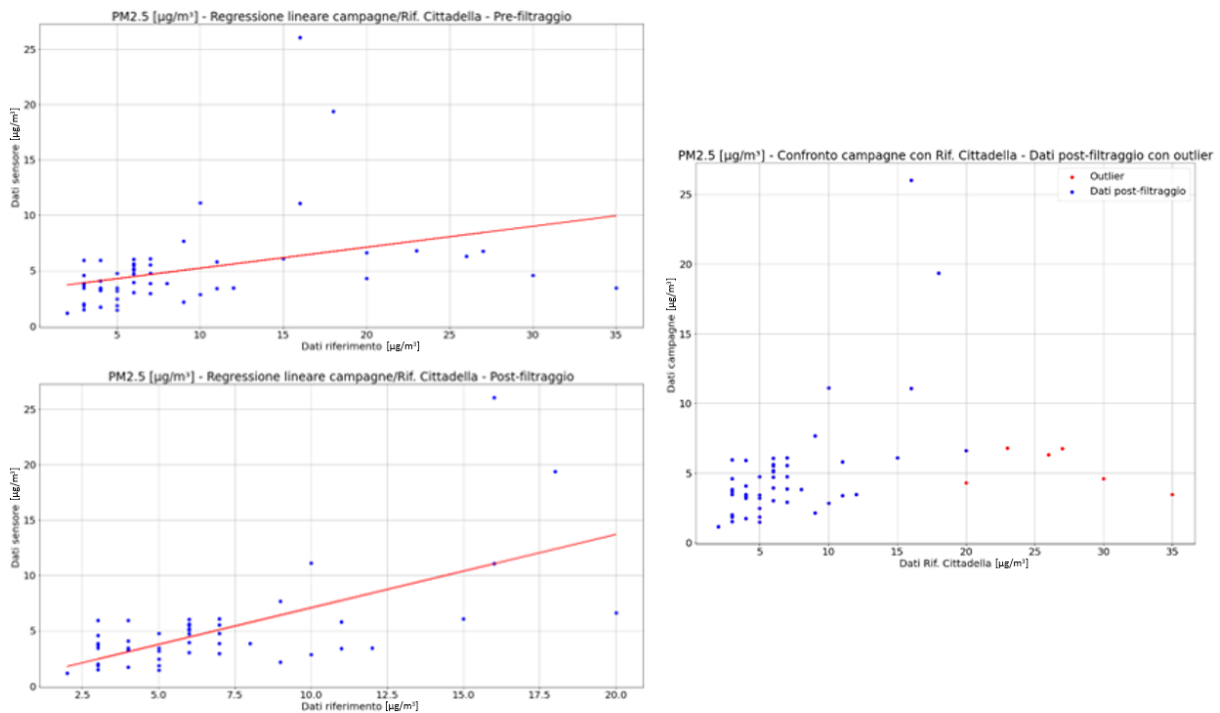


Figura 40: Grafici dell'analisi di regressione tra campagne da terra e riferimento di Cittadella; pre e post filtraggio

Nella Tabella [17] sono presentati i risultati complessivi del confronto tra le campagne di misura da terra e il riferimento di Cittadella.

PM2.5 [µg/m ³] - CONFRONTO STAZIONE DI RIFERIMENTO E CAMPAGNE	
Pendenza retta di regressione	0.66
Intercetta retta di regressione (Bias)	0.44
µ residui	1.89
σ residui	7.11
Coefficiente di Pearson	0.65
µ misure del riferimento	6.90
σ misure del riferimento	8.44
µ misure delle campagne	5.01
σ misure delle campagne	8.58

Tabella 17: Risultati dell'analisi di incertezza; confronto tra sensore fisso e stazione di riferimento

Osservando i valori in Tabella [16], si nota come c'è una migliore correlazione lineare tra i dati delle campagne e quelli del riferimento (coefficiente di Pearson pari a 0.65) rispetto al caso dei sensori fissi. Se si aggiunge a questa considerazione il fatto che la pendenza della retta di regressione sia pari a 0.66, un bias di 0.44 e la pressoché uguaglianza tra le deviazioni standard delle misure, si può concludere che c'è una buona corrispondenza tra i due dataset. Le conclusioni appena tratte non possono considerarsi definitive, in quanto come si nota dalla Figura [39] la quantità di dati raccolti durante le campagne è nettamente minore. Tuttavia i valori appena descritti e la similarità di andamento possono dimostrare come l'utente non vada a influenzare in maniera significativa nella misura.

5.1.2 Risultati dell'analisi geografica

In Figura [41] si presenta la divisione in cluster in base alle coordinate GPS del dataset delle campagne, a confronto con la *heatmap*.

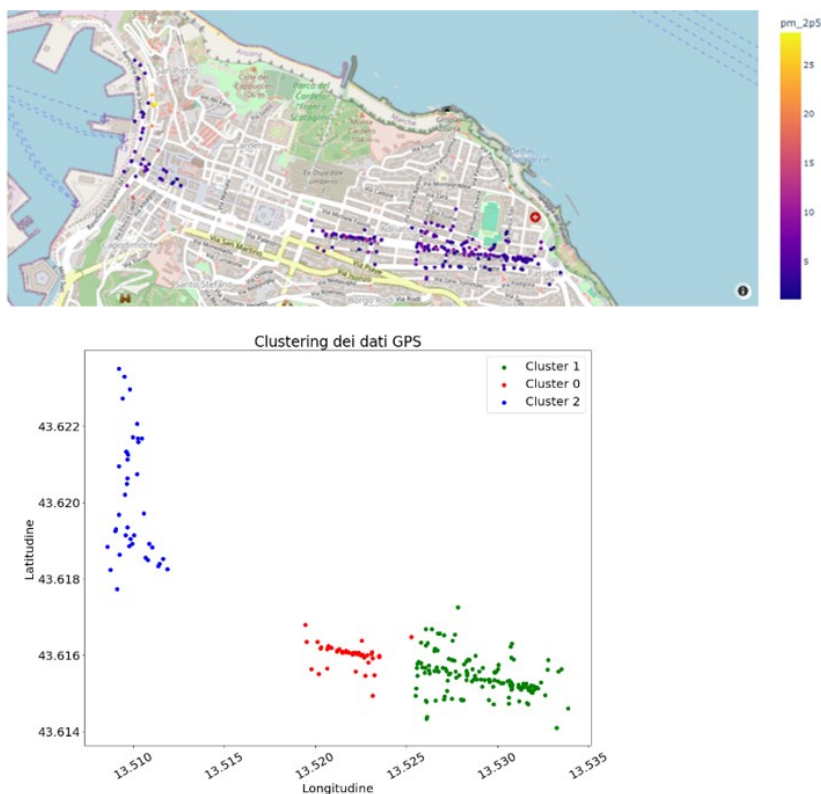


Figura 41: Divisione in cluster del dataset delle campagne da terra

Si sottolinea come la scelta di dividere in 3 cluster sia stata effettuata una volta visualizzata la *heatmap*.

Di seguito si presentano nella Tabella [18] le distanze calcolate durante l'analisi, con evidenziate in rosso le due che permettono di discutere l'influenza della distanza sulle misure. Nella Tabella [19] sono invece presentate la media e la STDV dei dataset e la loro differenza; anche qui sono evidenziati in rosso i valori di interesse, collegati alle distanze sopra citate.

	μ	σ
CAMPAGNE	5.2	3.9
Cluster 1	5.1	2.7
Cluster 2	4.7	2.2
Cluster 3	7.4	7.6
RIF. CITTADELLA	11.9	8.9
RIF. STAZIONE	13.4	9.6

Tabella 18: Media e deviazione standard dei cluster e dei riferimenti

DISTANZE IN METRI		DIFFERENZA	
Cluster 1-2	580.63	0.4	0.5
Cluster 1-3	1060.65	-2.3	-4.9
Cluster 2-3	1620.68	-2.7	-5.3
Cluster 1-RIF. CITTADELLA	975.48	-6.9	-6.3
Cluster 1-RIF. STAZIONE	2633.26	-8.3	-6.9
Cluster 2-RIF. CITTADELLA	1504.46	-7.3	-6.7
Cluster 2-RIF. STAZIONE	3143.41	-8.7	-7.4
Cluster 3-RIF. CITTADELLA	858.31	-4.6	-1.4
Cluster 3-RIF. STAZIONE	2121.39	-6.0	-2.0

Tabella 19: Risultati analisi geografica di Ancona

I valori segnati in rosso evidenziano come non sia possibile trovare una correlazione tra distanza e misura, dimostrando una stretta dipendenza tra posizione spaziale e valore misurato: si nota infatti come nonostante la distanza dei centroidi del cluster 1 e 3 dal riferimento di Cittadella sia pressoché la stessa, le due medie differiscono dal riferimento rispettivamente di 6.9 e 4.6 [$\mu\text{g}/\text{m}^3$] mentre le STDV di 6.3 e 1.4 [$\mu\text{g}/\text{m}^3$]. Queste differenze nelle misure possono trovare spiegazione nella distanza che separa i due centroidi, che è di 1060 [m].

5.1.3 Risultati dell'analisi sulle campagne di misura con drone

Come già spiegato nella Sezione 4.2.3.3, l'analisi delle campagne con drone è stata limitata per via della non comparabilità temporale delle misure, come dimostrato in Figura [42], dove vengono mostrati gli andamenti delle misure con drone e delle campagne ausiliarie da terra.

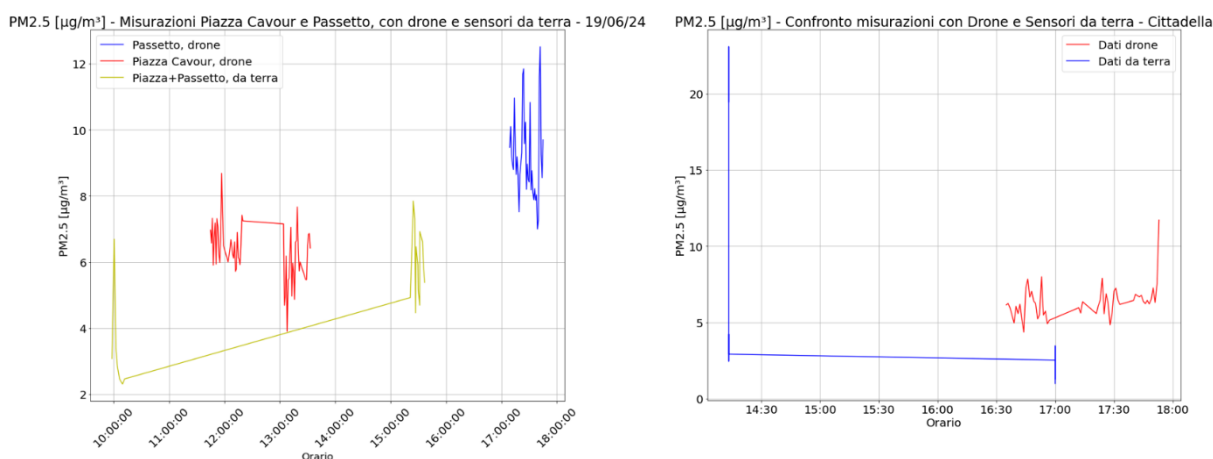


Figura 42: Confronto temporale tra campagne con drone e campagne ausiliarie da terra

Si presentano ora in Figura [43] i grafici che legano il dato misurato all'indice temporale e all'altezza sul livello del mare, mostrando come l'altitudine influisca notevolmente sulla misura, in quanto valori molto vicini temporalmente e spazialmente presentano valori differenti tra loro: sono presenti differenze anche di 2 [$\mu\text{g}/\text{m}^3$] su valori dell'ordine di 6-8 [$\mu\text{g}/\text{m}^3$] con ad

altezze distanti circa 15 [m] – a Piazza Cavour verso le 12 – mentre a Cittadella le arrivano anche a circa 4 $\mu\text{g}/\text{m}^3$ per misure a circa 60 [m] distanti in altitudine.

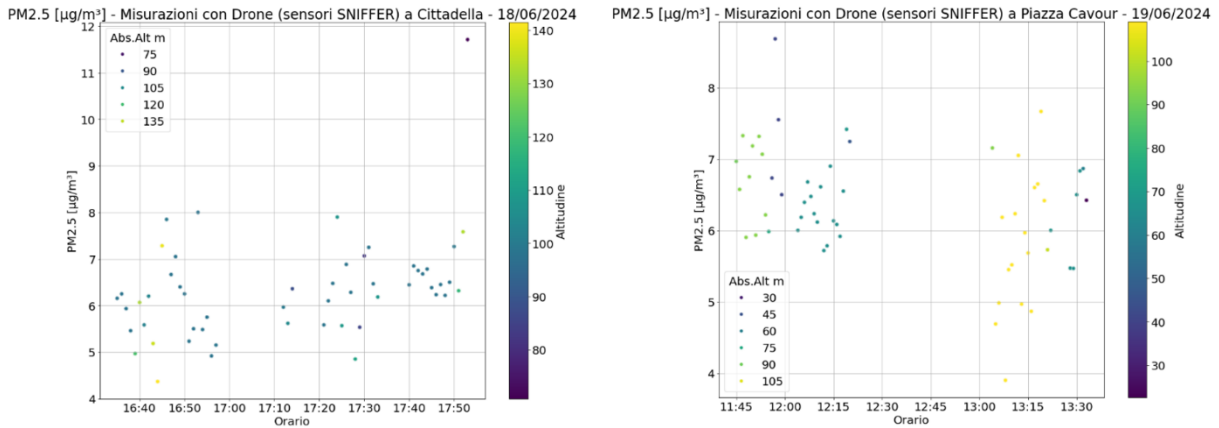


Figura 43: Grafici delle misure con drone in funzione dell'indice temporale e dell'altitudine sul livello del mare

5.2 Maroussi

Per la città pilota di Maroussi vengono presentati in questa Sezione i risultati della sola analisi completa che è stato possibile svolgere, descritta nella Sezione 4.2.4.1.

In Figura [44] è visualizzabile la divisione in 3 cluster dei 2 giorni selezionati, con annesse le heatmap per il confronto diretto.

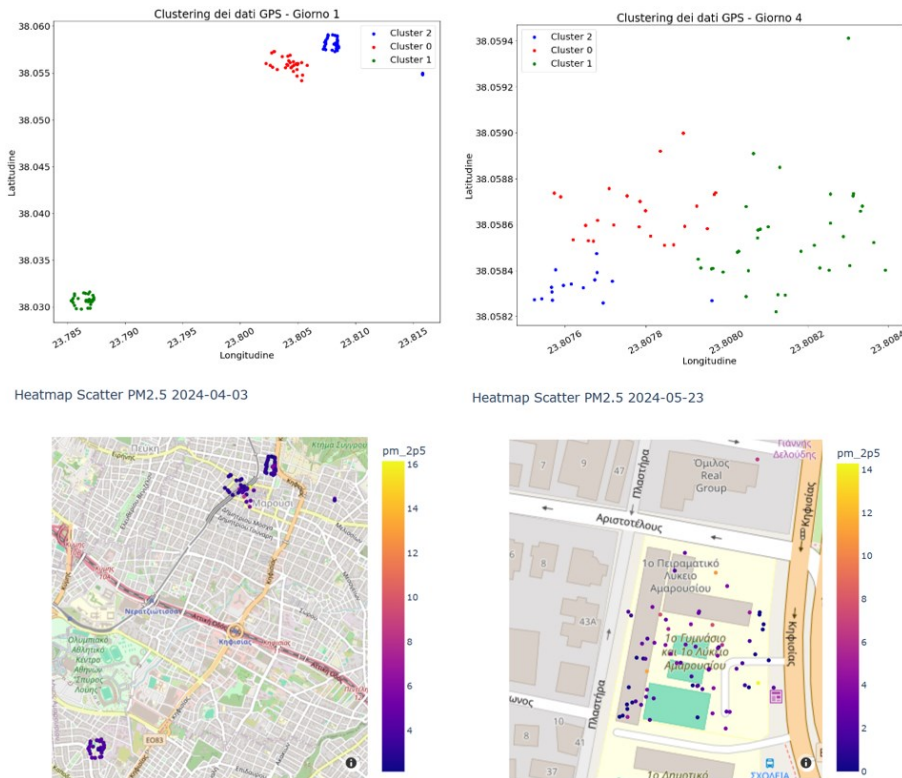


Figura 44: Divisione in cluster dei giorni selezionati

Si illustrano ora nella Tabella [20] le medie e deviazioni standard dei cluster di tutti e 10 i giorni con più misure, nella quale sono evidenziati i due giorni selezionati.

PM2.5 [$\mu\text{g}/\text{m}^3$]	03/04/2024	14/06/2023	28/06/2024	23/05/2024	02/04/2024	25/08/2023	16/06/2023	13/06/2024	19/06/2023	14/06/2024
GIORNO	1	2	3	4	5	6	7	8	9	10
MISURE	119	111	90	72	59	56	42	40	40	34
μ Cluster 1	4.1	7.8	5.0	2.8	9.6	6.2	2.0	26.0	9.0	21.3
σ Cluster 1	2.6	1.2	3.1	1.6	7.7	0.9	0.0	5.2	0.3	7.8
μ Cluster 2	4.1	8.2	8.0	2.9	4.1	6.6	3.9	24.8	9.3	24.5
σ Cluster 2	0.6	1.7	11.7	3.2	3.8	0.5	0.5	4.4	0.4	9.7
μ Cluster 3	3.3	7.9	5.3	2.0	13.4	6.6	4.8	30.0	9.7	20.5
σ Cluster 3	0.8	1.9	1.9	3.0	7.2	0.7	3.8	5.6	1.1	10.7
μ GIORNO	3.8	7.9	5.8	2.7	9.2	6.4	3.3	26.6	9.3	22.1
μ GIORNO RIF.1	3.4	12.7	9.3	3.7	12.0	21.8	12.1	8.9	20.9	6.7
μ GIORNO RIF.2	3.7	no data	9.8	no data	12.2	no data	no data	10.1	no data	8.7
DISTANZA MEDIA [m]	129.8	112.9	51.5	32.8	91.9	19.7	254.6	419.8	101.9	379.5

Tabella 20: Media e deviazione standard dei cluster dei 10 giorni con più misure

Per i due giorni evidenziati, nella Tabella [21] sono mostrate: in blu, le differenze tra le medie dei cluster, la media giornaliera e dei riferimenti; in verde, le distanze spaziale tra i punti geolocalizzati. I valori scritti con carattere rosso sono i più esemplificativi per comprendere l'influenza dell'ubicazione nella misura.

	GIORNO 1		GIORNO 4	
	DISTANZA	DIFFERENZA μ	DISTANZA	DIFFERENZA μ
Cluster 1-2	3198.1	0.1	35.1	-0.1
Cluster 2-3	3584.5	0.7	51.0	0.9
Cluster 1-3	439.5	0.8	38.3	0.8
Cluster 1-Rif.1	1054.4	0.7	1143.2	-0.9
Cluster 2-Rif.1	3639.9	0.7	1178.3	-0.8
Cluster 3-Rif.1	1235.7	-0.1	1146.8	-1.7
Cluster 1-Rif.2	3096.5	0.4	3453.8	X
Cluster 2-Rif.2	1083.7	0.4	3449.5	X
Cluster 3-Rif.2	3395.5	-0.4	3415.7	X

Tabella 21: Risultati analisi geografica di Maroussi

Si prenda il giorno 1. Si nota come non ci sia correlazione tra distanza spaziale e valori misurati: guardando ad esempio differenza tra le medie del cluster 1 e 2 (0.1 [$\mu\text{g}/\text{m}^3$]) e la distanza che li separa (3198 [m]) e la differenza tra le medie del cluster 1 e 3 (0.8 [$\mu\text{g}/\text{m}^3$]) e la distanza (439 [m]) si osserva un trend antitetico, ovvero che all'aumentare della distanza diminuisca la differenza. Prendendo invece le differenze tra il cluster 1 e 2 con il riferimento 1, notiamo che esse hanno lo stesso valore nonostante siano a distanze differenti (rispettivamente 1054 e 3639 [m]).

Si prenda ora il giorno 4. Qua è possibile notare come anche l'estrema vicinanza non porti come conseguenza la similitudine di valori: osservando le differenze tra le medie del cluster 1 e 2 e le differenze tra cluster 1 e 3 – con le distanze che non superano i 40 [m] – esse siano

rispettivamente di 0.1 e 0.8 $[\mu\text{g}/\text{m}^3]$. Questo trend è visualizzabile anche nel confronto dei cluster 1 e 3 con il riferimento 1, in quanto pur essendo praticamente alla stessa distanza da quest'ultimo (circa 1145 [m]) le differenze sono rispettivamente di 0.9 e 1.7 $[\mu\text{g}/\text{m}^3]$.

5.3 Saragozza

Per la città pilota di Saragozza i risultati presentati si limitano a mostrare come il filtraggio dei dati sia stato efficace, in quanto non è stato possibile svolgere ulteriori analisi per mancanza o incompatibilità temporale dei dati. In Figura [45] è mostrato il grafico rappresentante l'andamento delle misure prima e dopo del filtraggio preliminare.

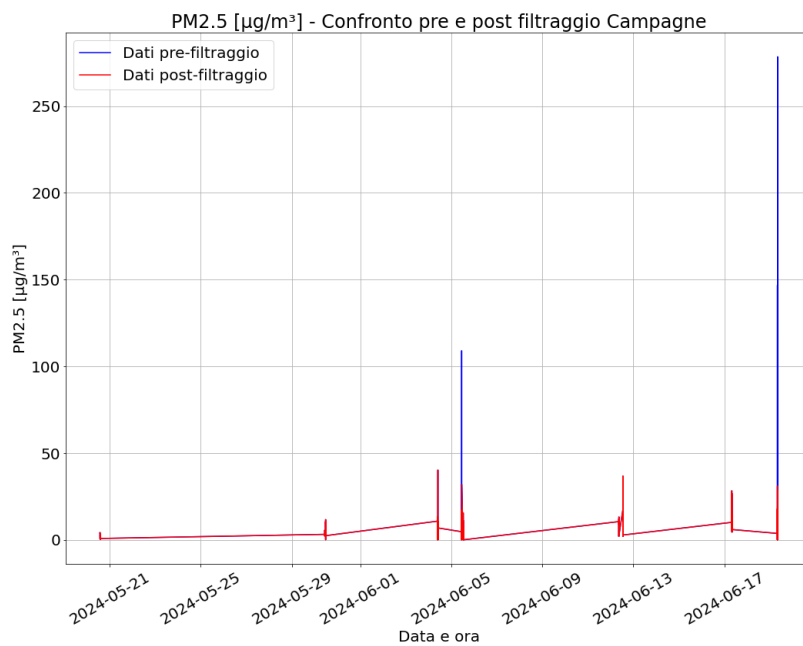


Figura 45: Grafico delle misure delle campagne da terra; pre e post filtraggio

Nella Tabella [22] sono rappresentati i valori dopo del filtraggio della media e deviazione standard del dataset delle campagne da terra.

CAMPAGNE DA TERRA		
μ	6.8	$[\mu\text{g}/\text{m}^3]$
σ	6	$[\mu\text{g}/\text{m}^3]$

Tabella 22: Media e deviazione standard delle campagne da terra di Saragozza

Capitolo 6: Conclusioni

Lo scopo dello studio è stato quello di valutare la qualità dei dati raccolti mediante utilizzo di sensori a basso costo (Low-Cost Sensors, LCSs) in un progetto di Citizen Science e indagare sull'influenza della distanza spaziale fra le misure. L'analisi eseguita nello studio ha permesso di verificare la qualità dei dati raccolti mediante l'utilizzo dei sensori a basso costo per la misura del particolato. Considerando innanzitutto il confronto tra i sensori fissi e la stazione di riferimento di Cittadella – poiché in questo caso le misurazioni sono prive di influenze dovute all'utente o a differenze spaziali tra rilevazioni successive: la deviazione standard di 5.90 [$\mu\text{g}/\text{m}^3$] contro 11.39 [$\mu\text{g}/\text{m}^3$] del riferimento dimostra come la dispersione delle misure sia quasi la metà; inoltre, si nota come l'andamento dei due dataset sia molto simile, con uno scostamento costante durante tutto il periodo in esame (scostamento che può essere in parte quantificato dalla media dei residui, pari a 5.33 [$\mu\text{g}/\text{m}^3$] e una bassa pendenza della retta di regressione, 0.27). La deviazione standard dei residui di 9.70 [$\mu\text{g}/\text{m}^3$] è da considerare invece come una prova dell'influenza della diversa collocazione dei due sistemi di misura in questione (circa 1 chilometro di distanza). Inoltre, facendo riferimento al confronto tra le campagne di misura da terra e il riferimento di Cittadella, si sottolinea come l'utente non influenzi la *performance* del sensore dato il coefficiente di Pearson pari a 0.65. Ciò permette di concludere che progetti di Citizen Science possono effettivamente essere di utilità nel monitoraggio della qualità dell'aria e nel *policy-making*. La seconda metodologia di analisi nello studio ha evidenziato come ci sia una stretta corrispondenza tra la posizione spaziale e la misura, in quanto si è sottolineato in più occasioni come la prima abbia una forte influenza sulla seconda. Si prenda ad esempio la città pilota di Maroussi: per giorni in cui le misure sono avvenute in un raggio di meno di 100 metri, le medie dei cluster differiscono tra loro di 0.8, 0.9 e -0.1 [$\mu\text{g}/\text{m}^3$] e dal riferimento (posto a circa 1150 [m]) di un ulteriore 0.8 [$\mu\text{g}/\text{m}^3$] (per un totale di 1.6, 1.7 e 0.8 [$\mu\text{g}/\text{m}^3$]). Questi dati dimostrano come la distanza tra le misure, anche non elevata, influisca nella misura stessa e mettono in luce la necessità di ampliare la capillarità del monitoraggio. I limiti dello studio sono da ricercare soprattutto nella metodologia di organizzazione del progetto e nella quantità di misure disponibili nei dataset, in quanto – sebbene abbia coinvolto un numero di utenti maggiore rispetto progetti simili precedenti e abbia svolto campagne più prolungate nel tempo – si è riscontrato come i dati raccolti sono stati in più circostanze non utilizzabili per l'analisi perché non compatibili tra loro. Futuri lavori che riprenderanno le metodologie utilizzate nel progetto e studio potranno porre attenzione sull'assicurare una maggiore copertura temporale delle misure, per permettere di conseguenza un'analisi sulla qualità dei dati più dettagliata. Inoltre, per meglio comprendere la magnitudo del contributo della posizione spaziale sulla misura, si evidenzia la possibilità di usare modelli differenti da quelli usati nello studio, come il modello di *Land Use Regression* (LUR). Grazie a quest'ultimo è infatti possibile indagare sull'influenza dell'ambiente circostante sul dato misurato (e.g. vicinanza a uno specchio d'acqua, a un parco, a un incrocio o a una strada trafficata), per distinguere le varie cause delle differenze tra misure distanti

spazialmente. Infine, un altro ambito interessante su cui basare future ricerche è quello delle misure con drone e la dipendenza che sussiste tra misura e altitudine.

Riferimenti bibliografici

- [1] Paul Gäbel, “Development of Air Quality Boxes Based on Low-Cost Sensor Technology for Ambient Air Quality Monitoring”, *sensors*, 2022.
- [2] Vanessa E. Alvear-Puertas, “Smart and Portable Air-Quality Monitoring IoT Low-Cost Devices in Ibarra City, Ecuador”, *sensors*, 2022.
- [3] Francesco Concas, “Low-Cost Outdoor Air Quality Monitoring and Sensor Calibration: A Survey and Critical Analysis”, *ACM Transactions on Sensor Networks*, 2021.
- [4] Sachit Mahajan, “Translating citizen-generated air quality data into evidence for shaping policy”, *Humanities & Social Sciences Communications*, 2022.
- [5] Pavel Buček, “Low-Cost Sensors for Air Quality Monitoring - The Current State Of The Technology and A Use Overview”, *Chemistry-Didactics-Ecology-Metrology*, 2021.
- [6] Aboubakr Benabbas, “Measure particulate matter by yourself: data-quality monitoring in a citizen science project”, *Journal of Sensors and Sensors Systems*, 2019.
- [7] Chris C. Lim, “Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea”, *Environment International*, 2019.
- [8] Sachit Mahajan, “A citizen science approach for enhancing public understanding of air pollution”, *Sustainable Cities and Society*, 2023.
- [9] Ye Kang, “Performance evaluation of low-cost air quality sensors: A review”, *Science of the Total Environment*, 2022.
- [10] Liliana Velea, “A community-Based Sensor Network for Monitoring the Air Quality in Urban Romania”, *Atmosphere*, 2023.
- [11] Pranvera Kortoçi, “Air pollution exposure monitoring using portable low-cost air quality sensors”, *Smart Health*, 2022.
- [12] Openpolis, “Le morti premature causate dal PM2.5”, *Ecologia e Innovazione*, 2022.
<https://www.openpolis.it/le-morti-premature-causate-dal-pm2-5/>
- [13] Ministero dell’Ambiente e della Tutela del Territorio e del Mare, “Concentrazione atmosferica degli inquinanti – Qualità dell’aria ambiente: Particolato (PM2,5)”.
https://www.salute.gov.it/imgs/C_17_paginaRelazione_1438_listaFile_itemName_2_file.pdf
- [14] Ministero della Salute, Opuscolo “Ozono (O3)”, 2015.
https://www.salute.gov.it/imgs/C_17_opuscoliPoster_283_ulterioriallegati_ulterioreallegato_8_alleg.pdf
- [15] Istituto Superiore della Sanità – Salute e ambiente, “Inquinamento: gli effetti dell’ozono sulla salute”. <https://www.epicentro.iss.it/ambiente/ozonoUe>
- [16] Ministero dell’Ambiente e della Tutela del Territorio e del Mare, “Concentrazione atmosferica degli inquinanti – Qualità dell’aria ambiente: Biossido di Azoto (NO2)”.
https://www.salute.gov.it/imgs/C_17_paginaRelazione_1438_listaFile_itemName_4_file.pdf

- [17] Openpolis, “È ancora elevata la concentrazione di No2 nell’aria”, *Ecologia e Innovazione*, 2022.
<https://www.openpolis.it/e-ancora-elevata-la-concentrazione-di-no2-nellaria/>
- [18] World Health Organization, “WHO/EURO:2022-3162-42920-63947 - Linee guida globali OMS sulla qualità dell'aria: particolato (PM2,5 e PM10), ozono, biossido di azoto, anidride solforosa e monossido di carbonio. Sintesi”, 2022.
- [19] Michael R. Giordano, “From low-cost sensors to high-quality data: A summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors”, *Journal of Aerosol Science*, 2021.
- [20] Settorezero, “Il sensore SDS011 e le polveri sottili”, 2020.
<https://www.settezero.com/wordpress/il-sensore-sds011-e-le-polveri-sottili/>
- [21] DigiKey, “Come utilizzare i sensori intelligenti di qualità dell'aria per il monitoraggio ambientale”, 2022. <https://www.digikey.it/it/articles/how-to-use-smart-air-quality-sensors-for-environmental-monitoring>
- [22] Talib Manshur, “A citizen science approach for air quality monitoring in a Kenyan informal development”, *City and Environment Interactions*, 2023.
- [23] Deliverable 5.7 progetto SOCIO-BEE – Definition & planning of pilots.R2
- [24] Deliverable 4. 8 progetto SOCIO-BEE – Wearables air quality hardware and software – 2 nd release
- [25] Deliverable 4.10 progetto SOCIO-BEE – AcadeMe web frontend and mobile app – 2 nd release