



UNIVERSITA' POLITECNICA DELLE MARCHE

FACOLTA' DI INGEGNERIA

Corso di Laurea triennale in Ingegneria Civile e Ambientale

Analisi statistica tra Covid-19 e inquinamento atmosferico

Statistical analysis between Covid-19 and air pollution

Tesi di Laurea di: **Marcozzi Mario** Relatore: **prof. Passerini Giorgio**

Correlatore: **ing. Virgili Simone**

A.A. 2019/ 2020

Sommario

1. Introduzione.....	3
1.1 Inquinamento atmosferico	4
1.2 Virus e pandemie	12
2. Materiali e metodi	21
2.1 Dati di inquinamento atmosferico.....	23
2.2 Dati di contagio e di mortalità del COVID-19	28
3. Risultati	34
3.1 Risultati per la regione Lombardia	36
3.2 Risultati per la regione Emilia Romagna	44
3.3 Riepilogo dei risultati ottenuti	53
4. Confronti e commenti	56
4.1 I coefficienti di Pearson a confronto	64
4.2 Influenza dell'inquinamento atmosferico a confronto con quella della densità abitativa sulla trasmissione del SARS-CoV-2	67
5. Conclusione	71
6. Bibliografia e sitografia	72

Prefazione

Con questa tesi si è deciso di analizzare se l'inquinamento atmosferico possa aver avuto o meno un'influenza sulla trasmissione e sulla mortalità del SARS-CoV-2 nelle provincie delle regioni di Lombardia ed Emilia Romagna. La scelta di queste due regioni è stata basata sul fatto che entrambe sono situate in Pianura Padana, una delle aree più inquinate di tutta l'Europa, in cui si verificano frequenti superamenti delle concentrazioni limite di legge di vari inquinanti dell'aria. Inoltre entrambe risultano due delle regioni più colpite dalla pandemia dovuta al COVID-19, quindi risultano tutte e due particolarmente adatte a questo tipo di analisi statistica.

In particolare, tra i possibili inquinanti dell'aria si è scelto di analizzare ozono troposferico e PM_{10} . Infatti, attualmente, essi sono due dei maggiori dell'aria perché entrambi risultano particolarmente dannosi per la salute della popolazione ed inoltre le loro concentrazioni superano spesso i limiti di legge a tutela della popolazione nelle regioni indagate. Si è invece escluso il $PM_{2,5}$ in quanto esso costituisce solo una frazione percentuale del PM_{10} .

Per l'ottenimento dei risultati è stato utilizzato il modello di regressione lineare semplice e le variabili indipendenti considerate, per le varie provincie, sono:

- Media PM_{10} del periodo 2016-2020;
- Numero di superamenti giornalieri di PM_{10} del valore di $20\mu g/m^3$ (nuova proposta dell'OMS come valore limite annuale), nel periodo 2016-2020;
- Numero di superamenti giornalieri di PM_{10} del valore di $50\mu g/m^3$ (limite di legge giornaliero), nel periodo 2016-2020;
- Numero di superamenti di O_3 del valore orario di $180\mu g/m^3$ (soglia di informazione), nel periodo 2016-2020.

Queste sono poi state elaborate, come presentato nel capitolo 3, insieme alle variabili dipendenti relative al SARS-CoV-2:

- casi totali al 28 maggio 2020;
- decessi/casi totali al 31 marzo 2020;
- decessi/casi totali al 30 aprile 2020.

Infine, per confrontare tutti i risultati ottenuti, si è deciso di utilizzare il coefficiente di Pearson che caratterizza tutte le rette di regressione lineare semplice che sono state ottenute a seguito di questa analisi statistica.

1. Introduzione

Il problema dell'inquinamento atmosferico sta assumendo negli ultimi anni un'importanza sempre più rilevante, soprattutto nelle aree urbane ed industriali dove si fa sentire la presenza dell'attività umana. C'è da sottolineare che nell'ultimo decennio in molti Paesi si è registrato un notevole miglioramento della qualità dell'aria grazie allo sviluppo di nuove politiche basate su fonti rinnovabili e non inquinanti. Tuttavia ancora in troppe aree del pianeta, Pianura Padana compresa, si registra una pessima qualità dell'aria con valori di inquinanti ben al di sopra dei limiti raccomandati che garantiscono la tutela della salute. Numerosi studi hanno provato l'esistenza di una correlazione tra concentrazione di inquinanti nell'aria e peggioramento delle condizioni di salute della popolazione. Proprio per questo motivo nasce anche l'esigenza di verificare l'esistenza di un'eventuale correlazione tra l'inquinamento atmosferico e l'infettività dei virus, in particolare quelli trasmissibili per via aerea.

Un altro grande problema per la salute umana è costituito dalle malattie virali, che da molti secoli sino ad oggi hanno messo in pericolo più volte la nostra specie. Dallo scorso dicembre 2019 l'argomento al centro di ogni attenzione è stata proprio la pandemia da COVID-19 causata dal Coronavirus 2 (SARS-CoV-2), un virus segnalato per la prima volta a Wuhan in Cina nel dicembre 2019 e successivamente diffusosi in tutto il mondo. Questo virus, trasmissibile principalmente per via aerea, ha provocato e ancora a luglio 2020 continua a provocare migliaia di decessi e centinaia di migliaia di nuovi casi infetti ogni giorno in tutto il mondo. La mancanza di una cura efficace contro il SARS-CoV-2 e la sua elevata infettività hanno reso necessaria l'adozione di una serie di misure di contenimento che hanno cambiato radicalmente le nostre abitudini, al fine di evitare l'eccessiva diffusione del virus. Di questo nuovo virus, sviluppatosi nello scorso 2019, conosciamo ben poco e ad oggi le incognite sono ancora tante: l'origine, i fattori che determinano la differente evoluzione dei contagi e della mortalità nelle diverse regioni e Paesi, o come il virus possa mutare ed evolversi... Quindi le domande aperte sono ancora molte e per questo risulta di fondamentale importanza lo studio e la comprensione di tutti quei fattori che possono aver contribuito alla formazione, diffusione ed evoluzione del SARS-CoV-2, anche allo scopo di non risultare più impreparati a un'emergenza di questo tipo che si potrebbe ripresentare in un prossimo futuro.

Con il lavoro che verrà di seguito presentato si è cercato di analizzare se l'inquinamento atmosferico abbia avuto o meno un'influenza sulla diffusione e sulla mortalità del

Coronavirus 2 in due delle regioni italiane più colpite: Lombardia ed Emilia Romagna. È bene precisare che questo studio non vuole imporsi come un risultato certo o assoluto, ma si è cercato trovare una possibile relazione tra alcuni degli inquinanti contenuti nell'aria e SARS-CoV-2, non ancora verificata sperimentalmente, anche grazie al supporto di altri studi effettuati riguardo questo argomento.

Prima di descrivere i metodi ed i risultati di quest'analisi è necessario introdurre nel dettaglio le due tematiche centrali.

1.1 Inquinamento atmosferico

L'inquinamento atmosferico può essere definito come l'insieme di tutti gli agenti fisici, chimici e biologici che modificano le caratteristiche naturali dell'atmosfera terrestre. L'inquinamento dell'aria è il risultato di tre fenomeni distinti:

- Emissione: emanazione delle sostanze nocive alla fonte;
- Trasmissione: diffusione e in parte trasformazione delle sostanze nocive nell'aria;
- Immissione: concentrazione o deposizione di inquinanti atmosferici nel luogo d'azione.

L'inquinamento atmosferico può essere suddiviso in due tipologie:

- Naturale: causato prevalentemente da eruzioni vulcaniche, processi biologici, formazione di polveri in seguito all'azione del vento.
- Di origine antropica: dovuto principalmente ad attività industriali, riscaldamento domestico e traffico veicolare.

A rendere questo fenomeno particolarmente pericoloso per la salute umana è il fatto che la maggior parte degli inquinanti rilasciati nell'aria non si disperde omogeneamente per tutta l'altezza dell'atmosfera terrestre, ma rimangono confinati all'interno della parte bassa della troposfera (strato dell'atmosfera più vicino alla superficie terrestre che raggiunge all'incirca i 10 km di altezza) nello strato limite planetario, anche detto Planetary Boundary Layer (PBL). Il PBL è definito come la parte di atmosfera che è direttamente influenzata dalla presenza della superficie terrestre attraverso lo scambio di calore e quantità di moto e che risponde alle forzanti termiche e meccaniche in un tempo non superiore all'ora. L'altezza del PBL, diversamente dagli altri strati dell'atmosfera, non è fissa ma può variare da poche centinaia di metri fino a circa 3000 metri dalla superficie terrestre. L'altezza limitata del PBL favorisce la concentrazione e il potenziale ristagno in aria di sostanze inquinanti. L'altezza dello strato limite planetario dipende

dalle condizioni meteorologiche e dalle caratteristiche orografiche del suolo e dall'ora del giorno; normalmente raggiunge il suo massimo nel primo pomeriggio, e va diminuendo nelle ore successive, fino a raggiungere il suo minimo nelle ore notturne. Quindi, in generale, le peggiori condizioni di inquinamento si verificano durante la notte/all'alba quando l'altezza del PBL è più ridotta, mentre durante la giornata la qualità dell'aria risulta migliore perché il PBL ha un'altezza maggiore.

I meccanismi principali con cui si può muovere una certa emissione in atmosfera sono prevalentemente tre:

- Avvezione: trasporto dell'emissione in una certa direzione dovuto al vento dominante.
- Galleggiamento: provoca la salita o la discesa dell'emissione. Per comprendere come avviene questo movimento è necessario introdurre il "Lapse Rate" (adiabatico secco), definito come il tasso di diminuzione standard della temperatura con la quota. Solitamente una data emissione in atmosfera all'inizio ha temperatura maggiore dell'aria circostante (e quindi densità minore) ed inizierà a salire raffreddandosi con il ritmo del Lapse Rate. In condizioni di atmosfera stabile (tipica della notte), la temperatura dell'aria diminuisce con la quota meno del Lapse Rate e quindi a un certo punto la particella d'aria considerata inizierà a scendere perché si raffredda più velocemente dell'aria circostante, tornando al suolo e facendo registrare livelli di inquinamento più elevati. In condizioni di atmosfera instabile (tipica del giorno), invece, la temperatura dell'aria diminuisce con la quota più velocemente del Lapse Rate e quindi una data emissione a temperatura maggiore dell'aria circostante tenderà a salire senza tornare mai al suolo perché la sua temperatura sarà sempre maggiore di quella dell'aria circostante; di conseguenza in questa situazione al suolo si avranno minori concentrazioni di inquinanti vicino al suolo.
- Turbolenza: provoca l'allargamento e la dispersione dell'emissione per via del rimescolamento dell'aria. Può essere distinta in turbolenza meccanica (dovuta all'azione del vento che urta con edifici ed oggetti), e turbolenza dovuta alle termiche (correnti di aria calda che si formano in seguito al riscaldamento del suolo da parte del sole e che salgono verso l'alto causando un rimescolamento per tutta l'altezza del PBL).

Ci sono poi anche altri meccanismi di trasporto e dispersione secondari dovuti alle brezze, e a fenomeni locali ed orografici.

Fatta questa breve panoramica sulla dispersione delle particelle in atmosfera, è possibile caratterizzare i vari inquinanti; questi si suddividono in:

- Inquinanti primari: sono quelli che vengono emessi dalle varie sorgenti direttamente in atmosfera, cioè non subiscono alcuna modifica una volta emessi nell'aria.
- Inquinanti secondari: sono quelli che hanno origine da reazioni chimico-fisiche tra gli inquinanti primari, o tra i primari e i componenti naturali dell'atmosfera.

Tra i principali inquinanti atmosferici, potenzialmente più pericolosi per la salute della popolazione, si possono citare gli ossidi di azoto (NO_x), l'Ozono troposferico (O_3), il particolato (in particolare PM_{10} e $\text{PM}_{2,5}$), metalli pesanti, benzene e Idrocarburi Policiclici Aromatici (IPA).

Tra quelli appena citati, in questo studio sono stati presi in considerazione PM_{10} ed Ozono troposferico perché estremamente pericolosi per la salute umana ed anche perché presenti nell'aria delle regioni indagate in elevate concentrazioni (spesso ben al di sopra dei limiti previsti dalla normativa).

Il PM_{10} è un inquinante sia primario che secondario. È definito come l'insieme delle particelle, solide o liquide, con diametro aerodinamico inferiore o uguale a $10\mu\text{m}$. Quindi una percentuale sul totale (tipicamente 50-70%) è costituita da $\text{PM}_{2,5}$, definito in modo analogo al PM_{10} come l'insieme delle particelle aventi diametro aerodinamico minore o uguale a $2,5\mu\text{m}$, e proprio per questo motivo non è stato considerato nell'analisi statistica.

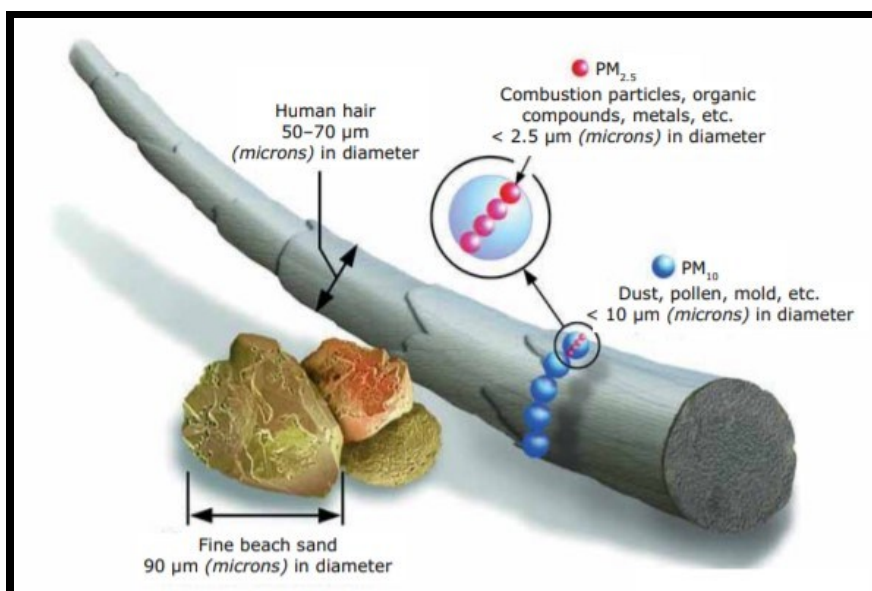


Figura 1: dimensioni del PM_{10} e $\text{PM}_{2,5}$ a confronto con quelle di un capello umano. Fonte: Environmental Protection Agency (EPA).

Il PM₁₀ può essere sia di origine antropica (derivante principalmente dalla combustione di biomasse, motori diesel e per riscaldamento degli edifici) o di origine naturale (dovuto prevalentemente ad incendi boschivi, erosione del suolo da parte del vento ed eruzioni vulcaniche). Secondo L'agenzia internazionale per la ricerca sul cancro (IARC) il PM₁₀ è stato classificato come cancerogeno gruppo 1, categoria riservata alle sostanze con sufficiente evidenza di cancerogenicità per l'uomo. Questo inquinante risulta particolarmente pericoloso perché rimane nell'aria per un tempo abbastanza lungo e può essere trasportato per lunghe distanze. Inoltre è caratterizzato da una composizione chimica molto variabile, anche in base alla stagione e alla posizione geografica (Perrino, 2020), poiché sulla sua superficie possono aderire altri inquinanti come gli IPA o i metalli (ad esempio piombo, nichel, cadmio, arsenico, vanadio, cromo), rendendolo ancora più nocivo per l'uomo. Il PM₁₀ è in grado di penetrare nel tratto respiratorio superiore (naso, faringe e trachea), tuttavia esso contiene anche il particolato fine con diametro aerodinamico inferiore a 2,5µm che è in riesce a penetrare profondamente nei polmoni e particolato ultra fine con dimensioni minori a 0,1µm che può penetrare fino agli alveoli polmonari e nel sistema cardiovascolare. Per cui gli effetti del PM₁₀ sulla salute dipendono molto dalla sua concentrazione dell'aria, ma anche dalla sua composizione chimico-fisica; tra i principali vi sono patologie acute e croniche a carico dell'apparato respiratorio (asma, bronchiti, enfisema, allergia, tumori) e cardio-circolatorio (aggravamento dei sintomi cardiaci nei soggetti predisposti). Le fasce della popolazione che possono risultare maggiormente interessate da questi effetti negativi sono anziani, bambini e persone sensibili (come gli asmatici e le persone con malattie polmonari e cardiache preesistenti). Stando a quanto riportato da numerosi studi, non è possibile individuare una soglia al di sotto della quale il PM₁₀ non costituisce un pericolo per la salute umana perché esso può risultare dannoso a qualsiasi concentrazione, per cui è raccomandato di mantenere una sua concentrazione nell'aria più bassa possibile. Le linee guida sulla qualità dell'aria o AQG (Air Quality Guidelines) dell'Organizzazione Mondiale della Sanità (OMS) hanno introdotto dei valori limite da rispettare allo scopo di limitare e controllare l'inquinamento nelle zone urbane, e per evitare l'eccessiva esposizione della salute pubblica a livelli di inquinamento eccessivamente elevati. Nel 2016, il 91% della popolazione mondiale viveva in luoghi in cui i livelli definiti dalle linee guida sulla qualità dell'aria dell'OMS non erano stati rispettati. Gli AQG hanno fissato per il PM₁₀ una media limite giornaliera di 50µg/m³, mentre la media annuale deve

risultare inferiore ai $20\mu\text{g}/\text{m}^3$. Inoltre, l'OMS ha stimato che riducendo la concentrazione media annuale di PM_{10} da $70\mu\text{g}/\text{m}^3$ a $20\mu\text{g}/\text{m}^3$ è possibile ridurre le morti premature correlate all'inquinamento atmosferico (nel 2016 stimate pari a circa 4,2 milioni in tutto il mondo) del 15% circa (World Health Organization, 2018).

A livello nazionale, i limiti per il PM_{10} sono stabiliti dal Decreto Legislativo 13 agosto 2010, n.155: “Attuazione della direttiva 2008/50/CE relativa alla qualità dell'aria ambiente e per un'aria più pulita in Europa”. Il decreto stabilisce valori limite per l'esposizione al PM_{10} riguardanti sia la concentrazione annua ($40\mu\text{g}/\text{m}^3$), sia quella giornaliera ($50\mu\text{g}/\text{m}^3$), da non superare più di 35 volte per anno civile. Il valore limite per la concentrazione annua di $40\mu\text{g}/\text{m}^3$ supera ampiamente il valore massimo annuale stabilito dall'OMS di $20\mu\text{g}/\text{m}^3$. Tuttavia i limiti definiti a livello nazionale vengono ancora superati nel 2016 in molte province italiane.



Figura 2: superamenti del valore limite giornaliero di PM_{10} nel 2016. Fonte: ISPRA.

L'altro inquinante oggetto di studio è l'ozono troposferico (O_3); quest'ultimo è diverso dall'ozono stratosferico che si trova all'interno della stratosfera (fascia di atmosfera che si estende indicativamente tra i 10 km e i 50 km di altezza dalla superficie terrestre) e costituisce un'importante barriera contro i raggi UV ad alta energia generati dal sole, potenzialmente cancerogeni. Invece l'ozono che si forma nella troposfera (fascia di atmosfera che si estende fino a circa 10 km dalla superficie terrestre) è un inquinante secondario che si forma per mezzo di complesse reazioni fotochimiche in presenza di inquinanti primari quali gli ossidi d'azoto (NO_x) e i composti organici volatili (COV). Poiché la formazione dell'ozono troposferico è legata anche alla disponibilità di fotoni, le concentrazioni più elevate si registrano nei mesi più caldi dell'anno e nelle ore di massimo irraggiamento solare. Le principali fonti di emissione dei composti precursori

dell'ozono (NO_x e COV) sono il trasporto su strada, il riscaldamento civile e la produzione di energia. L'ozono troposferico può causare seri problemi alla salute dell'uomo: essendo un forte ossidante, è in grado di attaccare i tessuti dell'apparato respiratorio anche a basse concentrazioni, provocando irritazione agli occhi e alla gola, tosse e riduzione della funzionalità polmonare. La maggior parte di questi effetti sono a breve termine e cessano con il cessare dell'esposizione ad elevati livelli di ozono, ma è noto che possano sussistere anche danni derivanti da ripetute esposizioni di breve durata, come l'accelerazione del naturale processo di invecchiamento della funzione polmonare e l'insorgenza di problemi di asma. I soggetti più a rischio sono bambini, anziani, persone con malattie respiratorie, ma anche soggetti sani che praticano attività fisica all'aperto. Il limite raccomandato dall'OMS per l'ozono troposferico, calcolato come media mobile su 8 ore, è di $100\mu\text{g}/\text{m}^3$.

Per l'ozono, il D. Lgs.155/2010 prevede una soglia di informazione e una di allarme. La soglia di informazione (fissata a $180\mu\text{g}/\text{m}^3$ come media oraria) viene definita come il livello oltre il quale sussiste un rischio per la salute umana in caso di esposizione di breve durata per alcuni gruppi particolarmente sensibili della popolazione nel suo complesso ed il cui raggiungimento impone di assicurare informazioni adeguate e tempestive. La soglia di allarme (pari a $240\mu\text{g}/\text{m}^3$, calcolata come media oraria) invece rappresenta il livello oltre il quale sussiste un rischio per la salute umana in caso di esposizione di breve durata per la popolazione nel suo complesso ed il cui raggiungimento impone di adottare provvedimenti immediati. Il D. Lgs.155/2010 definisce per l'ozono anche un valore obiettivo per la protezione della salute, da valutarsi sulla base delle rilevazioni condotte sugli ultimi 3 anni di misura. In particolare, il valore obiettivo è pari a $120\mu\text{g}/\text{m}^3$ come massima media mobile giornaliera calcolata su otto ore. Il valore obiettivo è da non superare per più di 25 giorni per anno civile, come media su tre anni. Analogamente al PM_{10} , il valore obiettivo per la protezione della salute umana fissato dal decreto (pari a $120\mu\text{g}/\text{m}^3$) risulta più alto di quello stabilito dall'OMS ($100\mu\text{g}/\text{m}^3$). Analogamente a quanto accade per il PM_{10} , al 2016 i limiti fissati dal decreto risultano superati in molte province del nostro Paese, come si può notare anche dalla seguente figura.



Figura 3: superamenti dell'obiettivo a lungo termine per l'ozono ($120 \mu\text{g}/\text{m}^3$ come media massima giornaliera calcolata su otto ore nell'arco di un anno civile). Fonte: ISPRA.

Ad aggravare ulteriormente la situazione italiana contribuisce il fatto che, in molte provincie e regioni, i limiti stabiliti dal D. Lgs.155/2010 sono troppo spesso ampiamente superati, sia per l'ozono che per il PM_{10} . Il caso emblematico è costituito dalla Pianura Padana, in cui si trovano anche entrambe le regioni oggetto di questo studio: Lombardia ed Emilia Romagna. La Pianura Padana si estende lungo l'Italia settentrionale ed è compresa principalmente entro il bacino idrografico del fiume Po delimitato dalle Alpi e Prealpi italiane a nord e ovest, dall'Appennino settentrionale a sud e dall'Alto Adriatico a est. Le conseguenze di questa particolare conformazione morfologica ed orografica

sono lo scarso ricambio di masse d'aria e basse velocità del vento, che causano frequenti fenomeni di inversione termica, intrappolando gli inquinanti emessi nell'aria vicino al suolo. Questo sfavorevole contesto geografico, unitamente ad una consolidata situazione di emissioni inquinanti dovute principalmente al traffico stradale ed al riscaldamento domestico, hanno contribuito a rendere quest'area una delle più inquinate di tutta l'Europa. Nonostante negli ultimi anni si sia verificata un'importante riduzione dei livelli di inquinamento in molte zone della Pianura Padana, come ad esempio nella Lombardia (Carugno, 2017), si registrano ancora concentrazioni di inquinanti spesso al di sopra dei limiti del D. Lgs.155/2010. Questo è ben visibile anche dalle precedenti figure 2 e 3 e da quella di seguito presentata.

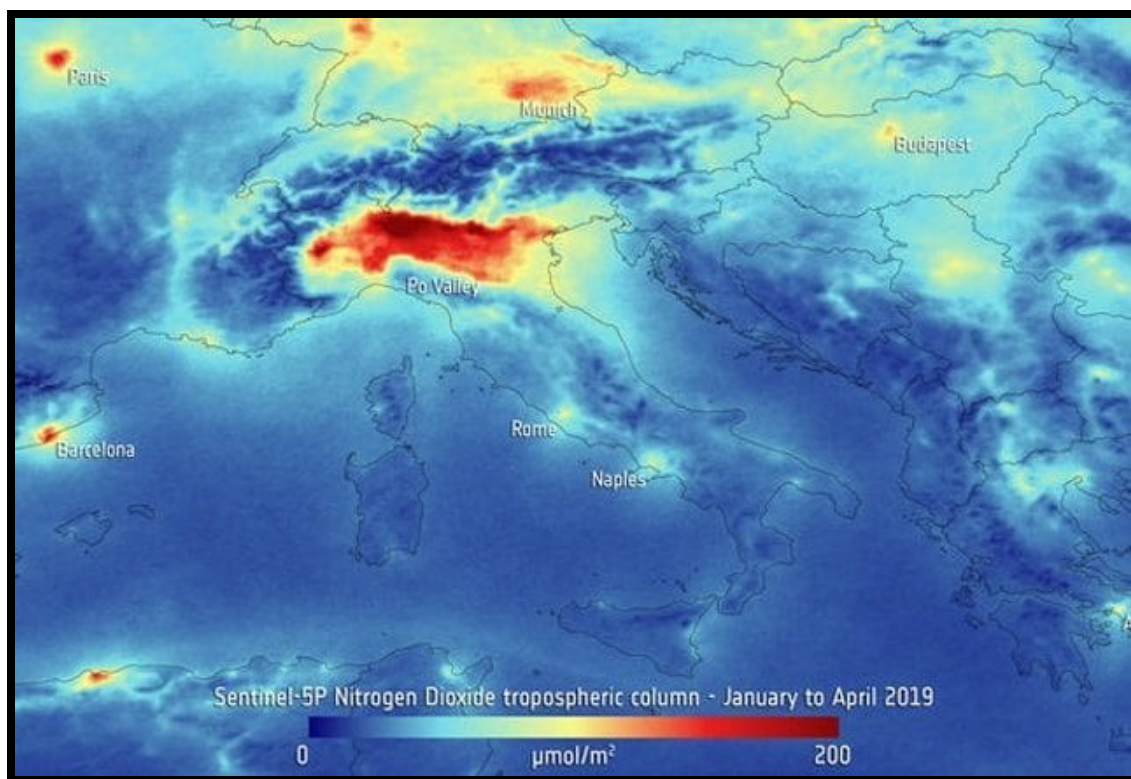


Figura 4: livelli di inquinamento da NO₂ rilevati dal satellite Sentinel-5P nei primi mesi del 2019. Fonte: European Space Agency (ESA).

A conferma di quanto appena detto, dalla figura soprariportata si nota chiaramente che la Pianura Padana risulta una delle aree più inquinate di tutto il continente europeo.

1.2 Virus e pandemie

Fin dai tempi più remoti l'uomo ha dovuto fronteggiare questa minaccia, invisibile ai nostri occhi, ma potenzialmente molto pericolosa. Il termine virus, coniato poco dopo la

loro scoperta alla fine del XIX secolo, è un termine latino che significa “veleno”, proprio allo scopo di indicarne la loro pericolosità per la nostra salute.

I virus sono dei microorganismi estremamente piccoli, solitamente con diametro compreso tra 20 e 300 nanometri e visibili solo al microscopio elettronico, costituiti da materiale genetico (DNA o RNA) racchiuso in un involucro di proteine (capside) e, in alcuni casi, anche in una membrana più esterna (detta pericapside) costituita da fosfolipidi e proteine. I virus non sono in grado di riprodursi (replicarsi) autonomamente, ma possono farlo esclusivamente all'interno delle cellule dei tessuti dell'organismo, causandone la distruzione o, per alcuni virus particolari, la trasformazione in cellule tumorali. La resistenza dei virus nell'ambiente è estremamente bassa, anche se alcuni di essi possono sopravvivere a lungo. I virus possono infettare tutte le forme di vita (persone, animali, piante, microrganismi ed anche altri virus) ed in particolare si trasmettono nell'uomo per via aerea, alimentare, attraverso rapporti sessuali o attraverso vettori (soprattutto insetti); possono causare disturbi locali a carico di diversi apparati o generali (sistemici) qualora si diffondano in tutto l'organismo. Alcuni tipi di virus possono essere combattuti con farmaci efficaci o prevenuti grazie alla vaccinazione specifica.

Quando si ha a che fare con una malattia infettiva particolarmente contagiosa, per la quale le persone non hanno immunità e capace di diffondersi rapidamente su scala globale si parla di pandemia. Una pandemia può avere un impatto estremamente negativo sulla salute della popolazione e sull'economia di tutto il pianeta. Questi effetti sono tanto più gravi quanto più i Paesi e le regioni inizialmente colpite dal virus che ha causato la pandemia, risultano impreparati a un'emergenza di questo tipo, non riuscendo a contenere la diffusione della malattia virale. A questo scopo, l'OMS ha realizzato un documento guida sulla preparazione a future pandemie nel 1999, di seguito aggiornato nel 2005 e 2009, che riporta una descrizione delle varie fasi pandemiche e le azioni da mettere in atto per contenere la pandemia durante ognuno dei diversi momenti. Le fasi descritte sono le seguenti:

- Fase 1: non si riportano casi di infezione tra gli esseri umani dovuti a virus influenzali di origine animale.
- Fase 2: un virus influenzale di origine animale, in circolazione tra animali domestici o selvatici, ha causato rari casi nell'uomo ed è quindi considerato una minaccia per una potenziale pandemia.

- Fase 3: un virus influenzale animale o riassortito umano-animale ha causato casi sporadici e piccoli cluster nella popolazione ma non ha provocato una trasmissione da uomo a uomo tale da alimentare focolai a livello di comunità.
- Fase 4: è stata accertata la trasmissione da uomo a uomo di un virus influenzale animale o riassortito umano-animale in grado di alimentare focolai a livello di comunità.
- Fase 5: lo stesso virus identificato ha causato estesi focolai a livello di comunità in due o più Paesi di una delle Regioni OMS.
- Fase 6: in aggiunta ai criteri definiti nella fase 5, lo stesso virus ha causato estesi focolai a livello di comunità in almeno un altro Paese di un'altra Regione OMS.
- Periodo successivo al picco: i livelli di influenza pandemica, nella maggior parte dei Paesi con sorveglianza adeguata, sono scesi dopo aver raggiunto il picco.
- Periodo post pandemico: i livelli di attività influenzale sono tornati a quelli tipici dell'influenza stagionale nella maggior parte dei Paesi con sorveglianza adeguata.

Tra le più recenti pandemie del passato si ricordano quelle del XX secolo. In questo secolo si sono verificate tre pandemie infettive, identificate in base alla presunta area di origine:

- 1) L'influenza spagnola (1918-1920): fu causata da un virus dell'influenza A, sottotipo H1N1, ed è stata l'epidemia più devastante della storia moderna. La sua caratteristica principale fu l'elevata mortalità, stimata al 2-3 %, che provocò un numero di morti compreso tra i 20 e più di 50 milioni in tutto il mondo (Johnson NPAS, 2002). Le cause di questa insolitamente alta mortalità non sono ancora del tutto chiare. L'aggressività dell'influenza spagnola sembra essere legata alle caratteristiche assunte dalle particelle virali in grado di suscitare un'eccessiva reazione del sistema immunitario non più efficacemente protettivo, ma partecipe del danno anatomico e quindi delle gravi conseguenze cliniche che portavano alla morte dei soggetti per polmonite. Quindi, paradossalmente, le persone più colpite dal virus furono quelle con il sistema immunitario più sano (ovvero giovani ed adulti con età inferiore ai 65 anni). Tuttavia non si esclude che la gravità memorabile e probabilmente unica della pandemia del 1918 potrebbe essere dipesa, almeno in parte, dalle condizioni di guerra e dalle infezioni batteriche secondarie in assenza di farmaci antimicrobici (Kilbourne, 2006).
- 2) L'influenza asiatica (1957-1958): anche questa fu causata da un virus dell'influenza A, sottotipo H2N2. Questa malattia virale fu responsabile di un numero

di decessi, principalmente per via di polmonite virale, stimato tra 1 e 2 milioni. Durante questa pandemia le persone più colpite furono quelle di età inferiore ai 70 anni. Il virus A/H2N2 non si dimostrò particolarmente infettivo e la sua mortalità fu sensibilmente più bassa della precedente influenza spagnola. Infine scomparì definitivamente nel 1968, quando mutò nel nuovo virus A/H3N2.

- 3) L'influenza di Hong Kong (1968-1969): fu causata dal virus H3N2 (anch'esso sottotipo dell'influenza A). Dal momento che il virus di Hong Kong differiva dal suo antecedente virus asiatico per il suo antigene HA, ma aveva mantenuto lo stesso antigene N2, è stato ipotizzato che il ridotto numero di decessi (stimato tra 750000 e 1 milione) e il suo ridotto impatto in alcune regioni del mondo rispetto alla precedente pandemia fosse imputabile a differenze nell'immunità acquisita nei confronti dell'antigene N2 (Monto, 1973) (Viboud, 2005). Per questo motivo l'influenza di Hong Kong è considerata la pandemia più delicata del XX secolo.

Anche nel XXI secolo si sono registrate delle pandemie: la prima risale al 2009, quando si è sviluppato un nuovo virus influenzale A (H1N1) che conteneva una combinazione unica di geni influenzali non precedentemente identificati in persone o animali. Questo nuovo virus, inizialmente denominato influenza suina per via della sua origine animale, è stato designato come virus pdm09 dell'influenza A (H1N1). Diversamente dalla tipica influenza stagionale, i soggetti più colpiti da questa malattia virale furono quasi esclusivamente al di sotto dei 65 anni. L'11 giugno 2009, l'OMS ha ufficialmente dichiarato l'esistenza di uno stato di pandemia da nuovo virus influenzale, con passaggio alla Fase 6 dei livelli di allerta pandemici individuati dal piano di preparazione e risposta alle pandemie influenzali. Il 10 agosto 2010 il direttore generale dell'OMS ha dichiarato conclusa la fase 6 dell'allerta pandemica e, pertanto, si è entrati nella fase post-pandemica. In questa fase, l'attività dell'influenza ritorna a livelli normali. Il CDC (Centers for Disease Control and Prevention) ha stimato che 151.700-575.400 persone in tutto il mondo sono morte a causa dell'infezione da virus pdm09 (H1N1) durante il primo anno di diffusione del virus e che circa l'80% dei decessi si è verificato in persone con età inferiore ai 65 anni.

Negli anni '60 fu scoperta una nuova famiglia di virus, i cosiddetti Coronavirus, che, negli ultimi decenni, sono stati i protagonisti di due diverse epidemie virali. I Coronavirus sono una vasta categoria di virus, di cui sette risultano patogeni anche per l'uomo; la maggior

parte di questi sono associati a lievi sintomi clinici, ad eccezione di due che negli ultimi anni hanno avuto conseguenze anche mortali sulla popolazione di molti Paesi:

- SARS-CoV: è responsabile della SARS (Severe Acute Respiratory Syndrome), una malattia respiratoria che ha avuto origine in Cina a fine 2002 e che nei casi più gravi portò alla morte per polmonite. Il SARS-CoV è un virus trasmissibile sia per via aerea che per contatto. Si ritiene che questo virus si sia inizialmente sviluppato nei pipistrelli, per poi essere trasmesso anche all'uomo (Li, 2005). Secondo le stime dell'OMS, tra il 2002 e il 2003, la SARS ha causato circa 800 decessi, con un tasso di mortalità di quasi il 10%. Dal 2004 in poi non si sono più registrati casi di infezione dovuti alla SARS.
- MERS-CoV: è il responsabile della malattia respiratoria MERS (Middle East Respiratory Syndrome) sviluppatasi nel 2012 in Arabia Saudita (de Groot, 2013). Secondo l'OMS ha causato, dalla sua comparsa fino al 2019, 858 decessi su un totale di 2494 casi di infezione. Quindi, il tasso di mortalità risulta di circa il 34%, considerevolmente più elevato rispetto alla SARS. Si pensa che anche questo Coronavirus abbia avuto origine dai pipistrelli, per poi essere trasmesso ai cammelli dromedari, che lo hanno infine trasmesso all'uomo (Mohd, 2016). Attualmente ancora si verificano casi sporadici di MERS.

Ai giorni d'oggi, nello scorso dicembre 2019, in Cina si è sviluppato un nuovo Coronavirus in grado di infettare anche l'uomo.

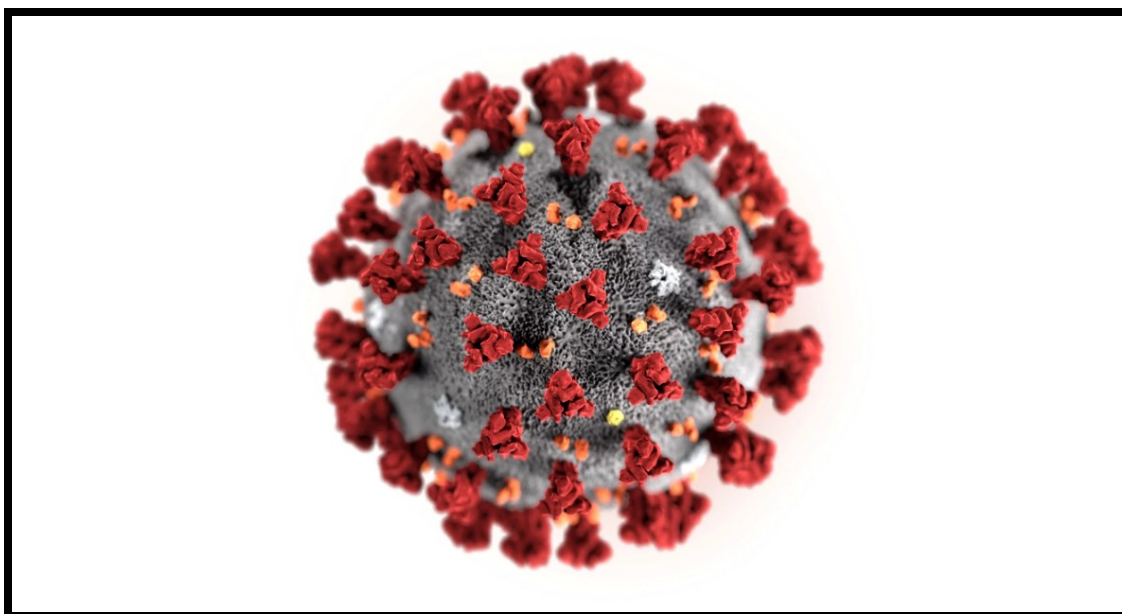


Figura 5: nuovo Coronavirus (SARS-CoV-2) al microscopio elettronico. Fonte: Centers for Disease Control and Prevention (CDC).

L'11 febbraio 2020, il direttore generale dell'OMS, il dott. Tedros Adhanom Ghebreyesus, ha annunciato che il nome della malattia causata da questo nuovo CoV era "COVID-19", che è l'acronimo di "coronavirus disease 2019". Questo nuovo virus è molto contagioso e si è rapidamente diffuso a livello globale. In una riunione del 30 gennaio 2020, secondo il regolamento sanitario internazionale (IHR, 2005), l'OMS ha dichiarato che il COVID-19 è un'emergenza sanitaria pubblica di interesse internazionale (PHEIC) in quanto si era diffusa in 18 paesi con quattro paesi che riportavano trasmissione all'uomo. Un'ulteriore data di riferimento è stata il 26 febbraio 2020, poiché il primo caso della malattia, non importato dalla Cina, è stato registrato negli Stati Uniti (USA). Successivamente l'OMS ha aumentato la minaccia all'epidemia di CoV a un livello "molto elevato", il 28 febbraio 2020. L'11 marzo, poiché il numero di casi COVID-19 al di fuori della Cina è aumentato di 13 volte e il numero di paesi coinvolti sono triplicati con oltre 118.000 casi in 114 paesi e oltre 4.000 decessi (Ghebreyesus, 2020), l'OMS ha dichiarato il COVID19 una pandemia.

Questo nuovo virus, inizialmente chiamato 2019-nCoV, è poi stato ufficialmente denominato come SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2) dall'ICTV (International Committee on Taxonomy of Viruses) per la sua estrema somiglianza al virus SARS-CoV responsabile della SARS nel 2002 (Zhou P. Y., 2020) (Zhu, 2020). L'origine del SARS-CoV-2 non è ancora del tutto certa, anche se numerosi

studi indipendenti ritengono che questo nuovo Coronavirus abbia avuto origine dai pipistrelli (Zhou P. Y., 2020) (Chan, 2020) (Wu, 2020) (Benvenuto, 2020) (Lu, 2020)... Le attuali evidenze suggeriscono che il SARS-CoV-2 si diffonde tra le persone, similmente al SARS-CoV, in modo diretto, in modo indiretto (attraverso oggetti o superfici contaminati) o per contatto stretto con persone infette attraverso secrezioni del naso e della bocca. Il periodo di incubazione del SARS-CoV-2 (periodo che intercorre tra l'esposizione all'agente virale e il manifestarsi dei sintomi della malattia) può variare da 2 a 14 giorni, mentre era compreso tra 2 e 10 giorni per il SARS-CoV. I sintomi della malattia sono molto simili a quelli della SARS: i più comuni sono febbre, tosse secca e spossatezza. Tuttavia lo spettro clinico di questa malattia può variare molto: da forme asintomatiche, a lievi sintomi semi-influenzali, a condizioni cliniche caratterizzate da insufficienza respiratoria che richiede ventilazione meccanica e supporto in terapia intensiva, a manifestazioni multi-organiche e sistemiche in termini di sepsi, shock settico e sindromi da disfunzione multipla di organi. Nei casi più gravi il SARS-CoV-2 può arrivare provocare la morte del paziente infetto, principalmente per insufficienza respiratoria o danni miocardici (Ruan, 2020).

I governi mondiali stanno lavorando per stabilire contromisure per cercare di contenere gli effetti devastanti del virus ed è stato stimato che i le rigide misure di contenimento adottate da alcuni Paesi europei potrebbero aver salvato 3 milioni di vite in 11 paesi (Flaxman, 2020). Le organizzazioni sanitarie stanno coordinando i flussi di informazioni ed emanando direttive e linee guida per mitigare al meglio l'impatto della minaccia. Allo stesso tempo, gli scienziati di tutto il mondo lavorano instancabilmente per cercare di capire a fondo lo spettro clinico della malattia e i meccanismi di trasmissione. Sebbene sia noto che COVID-19 sia causato dalla trasmissione da uomo a uomo, non è ancora chiaro se gli inquinanti atmosferici ambientali possano promuoverne la trasmissione. Prove sufficienti hanno strettamente collegato l'inquinamento dell'aria al verificarsi di numerose malattie respiratorie. Inoltre l'inquinamento atmosferico è anche associato alla trasmissione di malattie infettive; ad esempio è stato provato che una peggiore qualità dell'aria ha incrementato la mortalità per SARS (Cui, 2003). Ancora rimangono anche molte altre incertezze sia sull'interazione virus-ospite che sull'evoluzione della pandemia, con specifico riferimento ai tempi in cui raggiungerà il suo picco e quando si possa verificare una seconda ondata nei Paesi (come l'Italia) in cui il numero dei nuovi casi positivi rimane sostanzialmente basso dopo una prima grande ondata iniziale.

Sebbene nel tempo il tasso di mortalità (numero totale di decessi per una data malattia in relazione al numero totale di pazienti) di COVID-19 sia stato significativamente inferiore a quello delle epidemie di SARS e MERS, la trasmissione del SARS-CoV-2 è molto più efficace di quello dei virus precedenti, con un numero totale di morti molto più elevato. Al 31 luglio 2020 si sono verificati oltre 17 milioni di casi totali e quasi 700 mila decessi in tutto il mondo (Coronavirus disease (COVID-19), 2020). In particolare, in Italia sono stati registrati un totale di quasi 250 mila casi totali e circa 35 mila morti dall'inizio dell'epidemia.

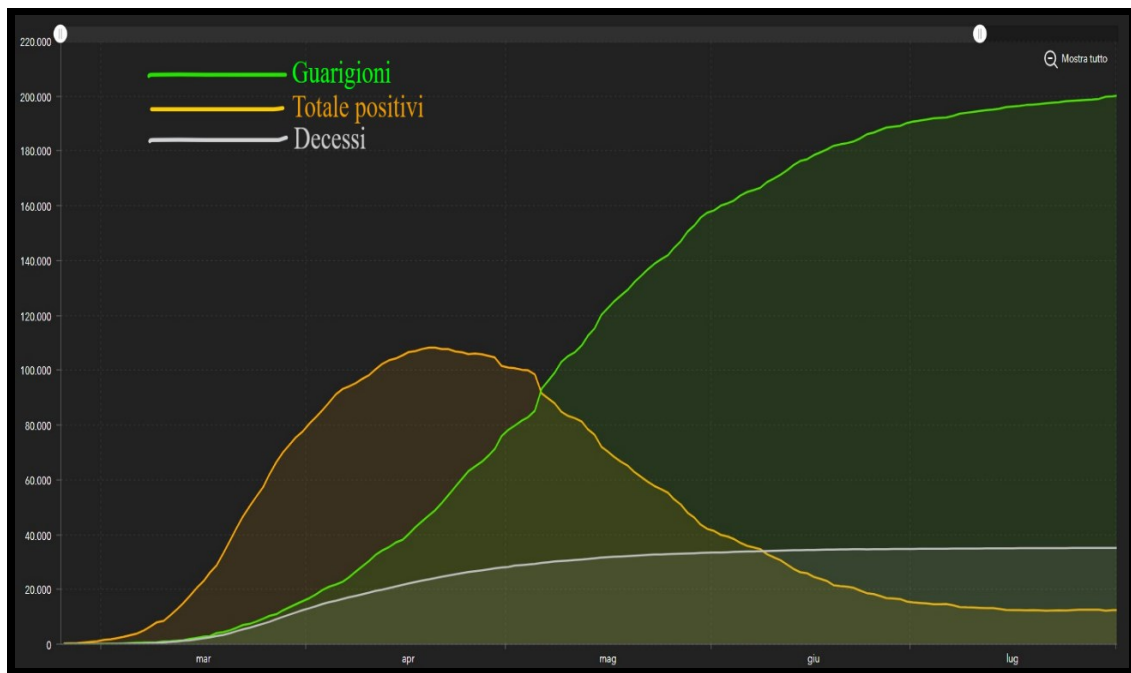


Figura 6: situazione in Italia fino al 31/07/2020. Fonte: Dipartimento della Protezione Civile e Ministero della Salute.

Al momento, le strategie terapeutiche per affrontare l'infezione sono solo di supporto e la prevenzione volta a ridurre la trasmissione nella comunità è la nostra arma migliore. Misure di isolamento restrittive in Cina hanno portato a una progressiva riduzione dei casi. Dalla Cina, la malattia si è poi diffusa anche in Europa. In Italia, inizialmente nelle regioni geografiche del nord e successivamente in tutta la penisola, le autorità politiche e sanitarie hanno compiuto sforzi incredibili per contenere un'onda d'urto che ha messo a dura prova il sistema sanitario nazionale. È grazie a loro, ma anche grazie agli sforzi compiuti dai cittadini italiani se ormai, da fine maggio a fine luglio 2020, si registra un numero di nuovi positivi relativamente basso ogni giorno, anche se non è ancora escluso che si possa registrare un nuovo aumento esponenziale dei casi in una seconda ondata.

Infatti al 31 luglio 2020 i nuovi casi sono ancora in aumento in molti Paesi, come si può vedere dalla seguente figura.

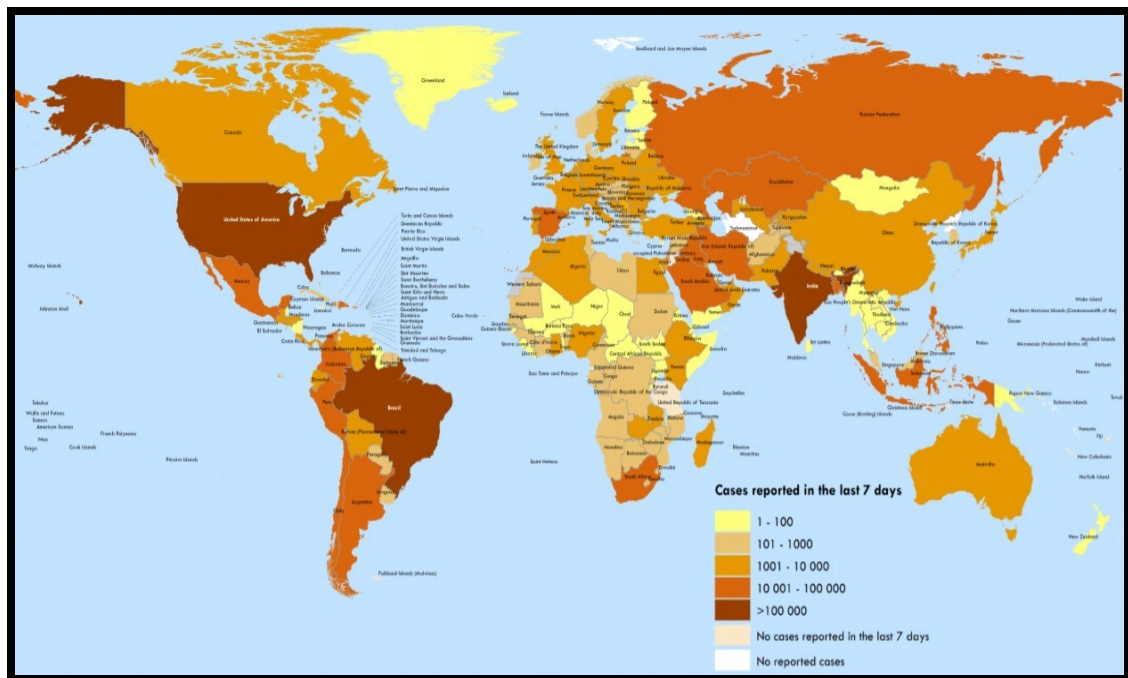


Figura 7: nuovi casi confermati per Regione dal 25 al 31 luglio 2020. Fonte: World Health Organization (WHO)

2. Materiali e metodi

Di seguito verranno descritte nel dettaglio tutte le fonti di dati che sono state utilizzate nel corso di questa analisi statistica ed i metodi che sono stati utilizzati nell'elaborazione dei dati per l'ottenimento dei risultati. È bene ribadire che le regioni oggetto di studio (Lombardia ed Emilia Romagna) sono state scelte perché entrambe sono situate nella Pianura Padana e quindi sono soggette ad elevati livelli di inquinamento dell'aria. Inoltre esse risultano essere due delle regioni più colpite dal SARS-CoV-2, quindi con un numero di dati sufficientemente elevato ed adatto all'analisi statistica. Infatti Lombardia ed Emilia Romagna sono rispettivamente al primo ed al terzo posto nella classifica regionale italiana per casi totali di COVID-19 al 31 luglio 2020. Secondo i dati riportati dalla Protezione Civile, Lombardia ed Emilia Romagna fanno registrare rispettivamente un totale di 96219 e 29670 casi totali fino al 31 luglio. Per quanto riguarda i decessi dovuti al COVID-19 fino a fine luglio, in Lombardia sono stati ben 16806 mentre in Emilia Romagna 4289. L'andamento nelle due regioni dei decessi e dei casi totali dovuti al SARS-CoV-2, dall'inizio dell'epidemia fino al 31 luglio, si può notare dalle seguenti figure.

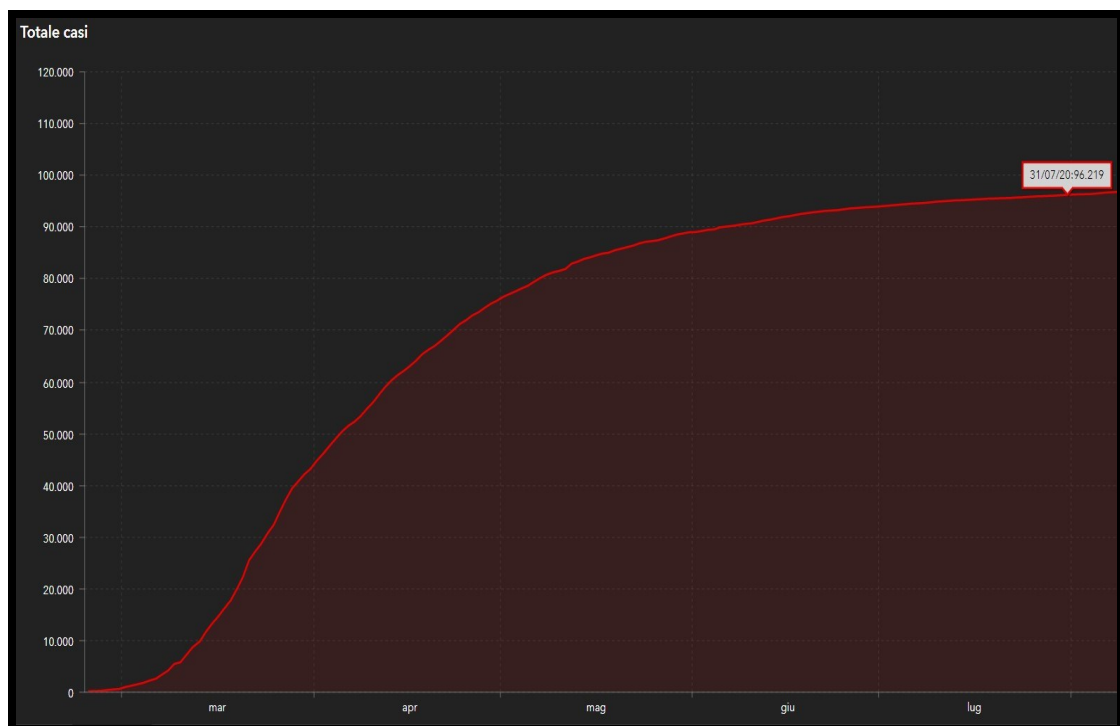


Figura 8: andamento dei casi totali in Lombardia fino al 31 luglio 2020. Fonte: Ministero della Salute e Dipartimento della Protezione Civile.

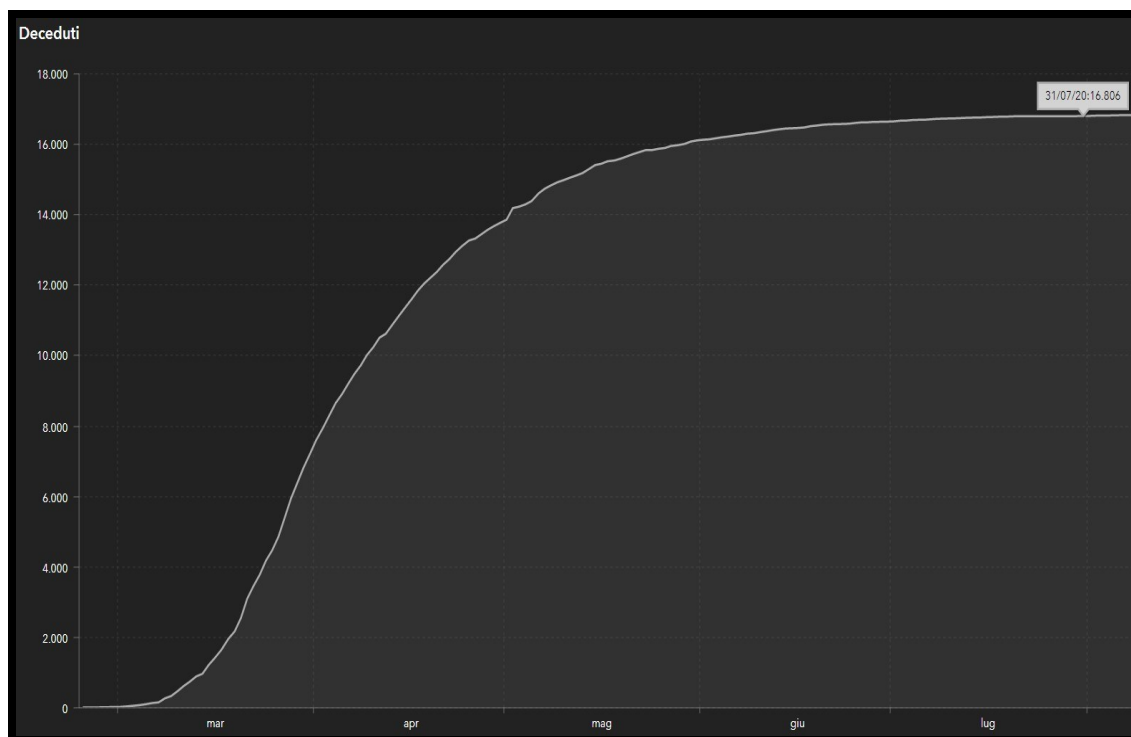


Figura 9: andamento dei decessi in Lombardia fino al 31 luglio 2020. Fonte: Ministero della Salute e Dipartimento della Protezione Civile.

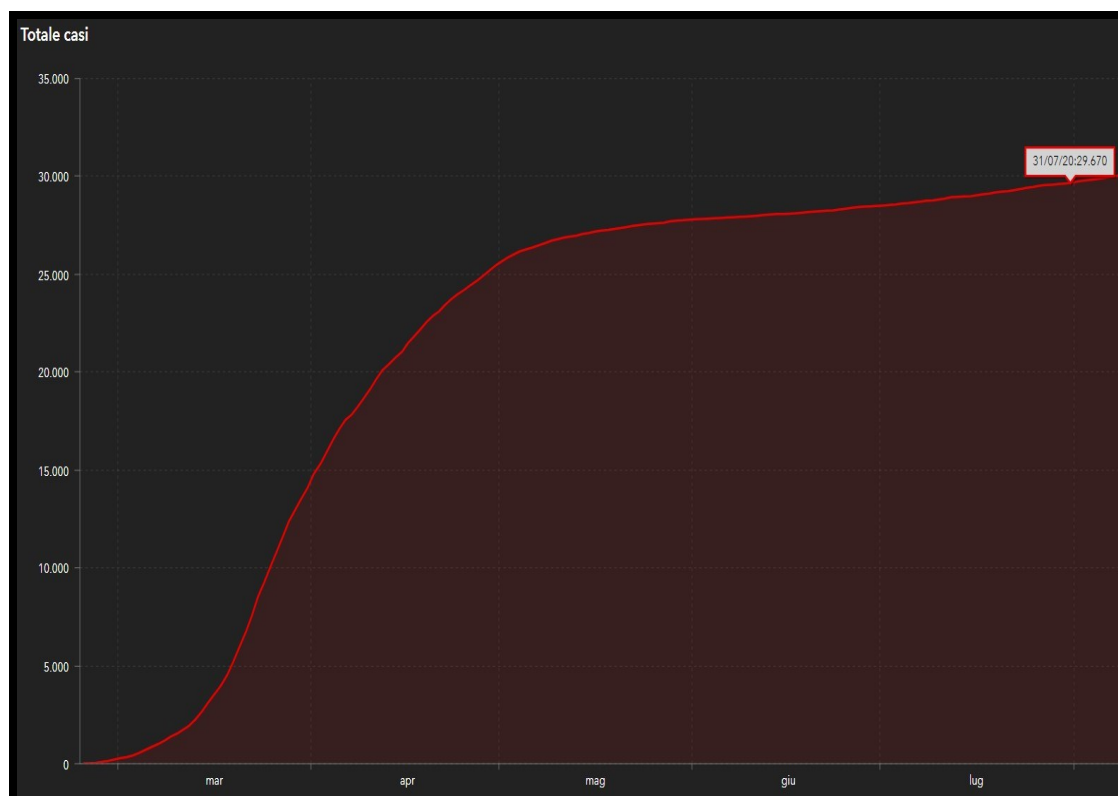


Figura 10: andamento dei casi totali in Emilia Romagna fino al 31 luglio 2020. Fonte: Ministero della Salute e Dipartimento della Protezione Civile.

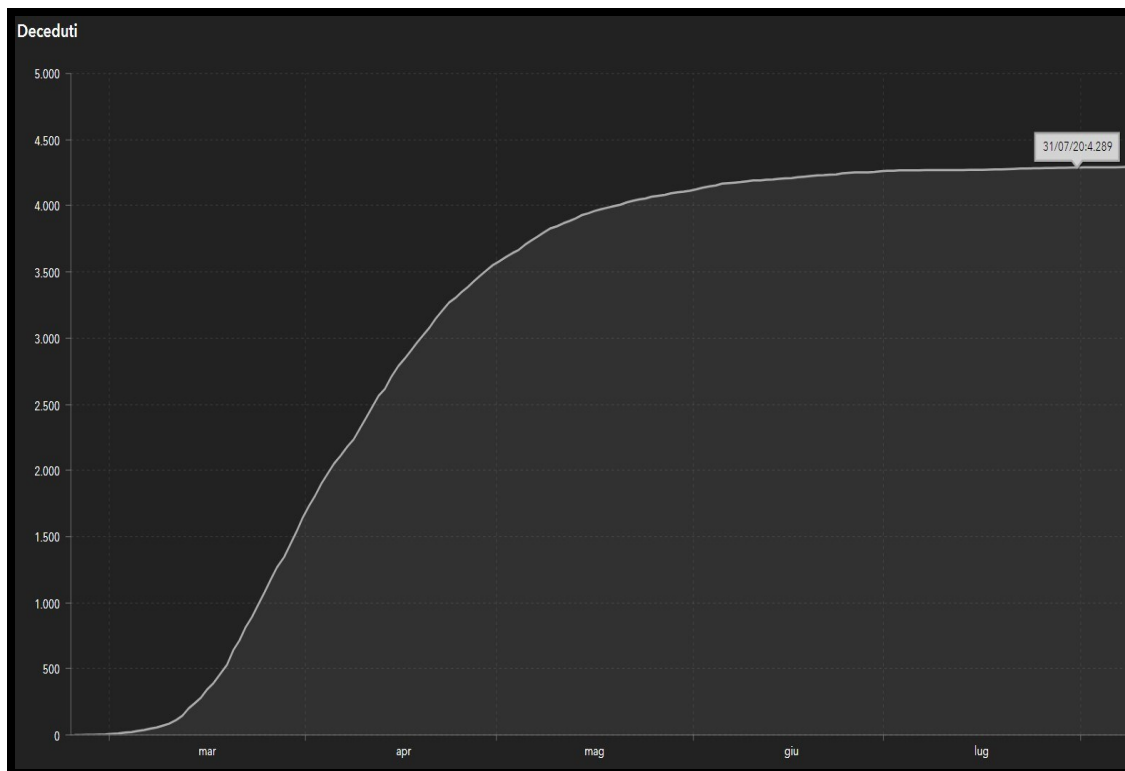


Figura 11: andamento dei decessi in Emilia Romagna fino al 31 luglio 2020. Fonte: Ministero della Salute e Dipartimento della Protezione Civile.

Le provincie oggetto dello studio sono: Bergamo, Brescia, Como, Cremona, Lecco, Lodi, Mantova, Milano, Monza, Pavia, Sondrio e Varese, per la Lomardia.

Invece quelle dell'Emilia Romagna sono: Bologna, Ferrara, Forlì-Cesena, Modena, Parma, Piacenza, Ravenna, Reggio Emilia e Rimini.

2.1 Dati di inquinamento atmosferico

Questi dati sono i primi che si sono resi necessari nell'analisi. I valori delle concentrazioni dei vari inquinanti sono misurate dalle regioni, in particolare dalla rete di centraline di monitoraggio dell'aria delle ARPA (Agenzie Regionali per la Protezione dell'Ambiente) disposte in tutto il territorio nazionale. Le ARPA si occupano del monitoraggio di tutti i temi ambientali (acqua, aria, suolo...) per la salvaguardia degli ecosistemi e della salute della popolazione in tutto il territorio italiano. In particolare, nelle due regioni indagate la misura e la pubblicazione dei suddetti dati spetta rispettivamente all'ARPA Lombardia ed all'ARPA Emilia Romagna. Come appena accennato, per la misura dei dati le varie ARPA fanno affidamento su numerose stazioni di rilevamento disposte in ogni regione e che solitamente misurano almeno i seguenti inquinanti: polveri sottili (PM₁₀ e PM_{2,5}),

ozono troposferico (O_3), ossidi di azoto (NO_x), biossido di zolfo (SO_2), monossido di carbonio (CO) e benzene (C_6H_6).

Il Decreto Legislativo n.155/2010 (che fissa le norme per la valutazione della qualità dell'aria) prevede che le regioni e le province autonome predispongano un programma per la misura della qualità dell'aria con stazioni fisse, coerente con le disposizioni introdotte dal decreto stesso. Il Ministero dell'ambiente, avvalendosi dell'ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale) e dell'ENEA (agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile), valuta la conformità del progetto alle prescrizioni normative.

Il numero delle stazioni di misurazione previste dal programma di valutazione deve essere individuato nel rispetto dei canoni di efficienza, efficacia ed economicità. Il numero di stazioni fisse dipende, in linea generale, dalla zonizzazione del territorio ed è funzione della popolazione di ciascuna zona e dello stato di qualità dell'aria: se le concentrazioni rilevate sono basse, sotto determinate soglie, il numero di punti da prevedere scende; viceversa per alte concentrazioni.



Figura 12: stazione di rilevamento. Fonte: ARPA Lombardia.

La rete di monitoraggio dell'ARPA Lombardia si compone di 85 stazioni fisse che monitorano costantemente, per mezzo di analizzatori automatici, la concentrazione dei

principali inquinanti in tutta la regione e forniscono dati in continuo con cadenza oraria. Inoltre si compone di 8 laboratori mobili per approfondire la situazione in specifiche località. Invece la rete di monitoraggio dell'ARPA Emilia Romagna si compone di 47 stazioni fisse. Non tutte le stazioni dispongono della stessa strumentazione, ma può cambiare in base al contesto in cui è posta ed agli inquinanti da misurare.

La legislazione italiana, costituita dal D. Lgs.155/10: "Attuazione della direttiva 2008/50/CE relativa alla qualità dell'aria ambiente e per un'aria più pulita in Europa", definisce che le regioni sono l'autorità competente nel campo della qualità dell'aria, e prevede la suddivisione del territorio in zone e agglomerati sui quali valutare il rispetto dei valori obiettivo e dei valori limite. La zonizzazione deve essere rivista almeno ogni 5 anni. Il D. Lgs.155/10 ha rivisto i criteri attraverso i quali realizzare la zonizzazione ai fini della valutazione della qualità dell'aria. In base ai criteri del suddetto decreto, la Lombardia ha distinto il territorio nelle seguenti zone:

- Agglomerati urbani: agglomerato di Milano, agglomerato di Bergamo, agglomerato di Brescia;
- Zona A: pianura ad elevata urbanizzazione;
- Zona B: zona di pianura;
- Zona C1: prealpi e appennino;
- Zona C2: montagna;
- Zona D: fondovalle.

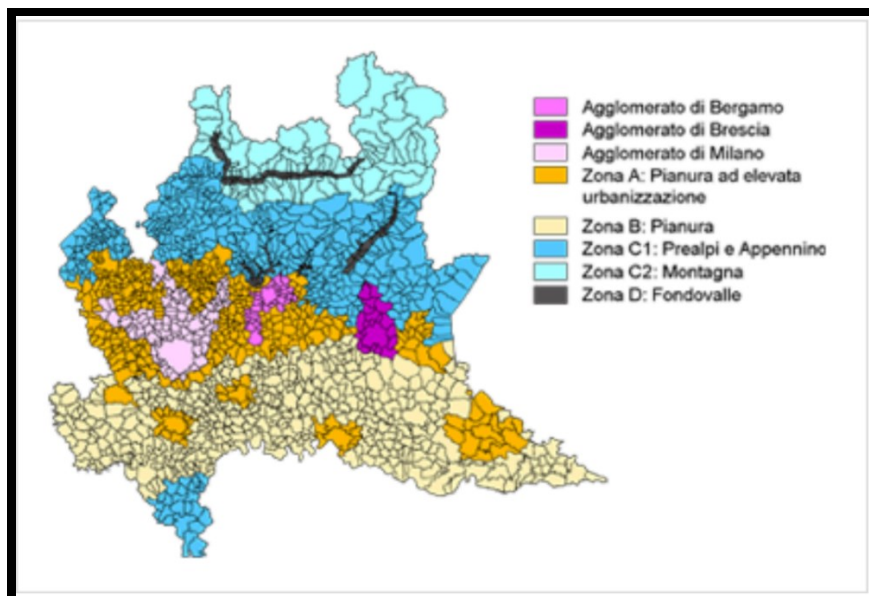


Figura 13: zonizzazione aggiornata della Lombardia. Fonte: ARPA Lombardia.

Invece l'Emilia Romagna ha suddiviso il territorio regionale nelle zone di seguito riportate:

- Agglomerato;
- Pianura est;
- Pianura ovest;
- Appennino.

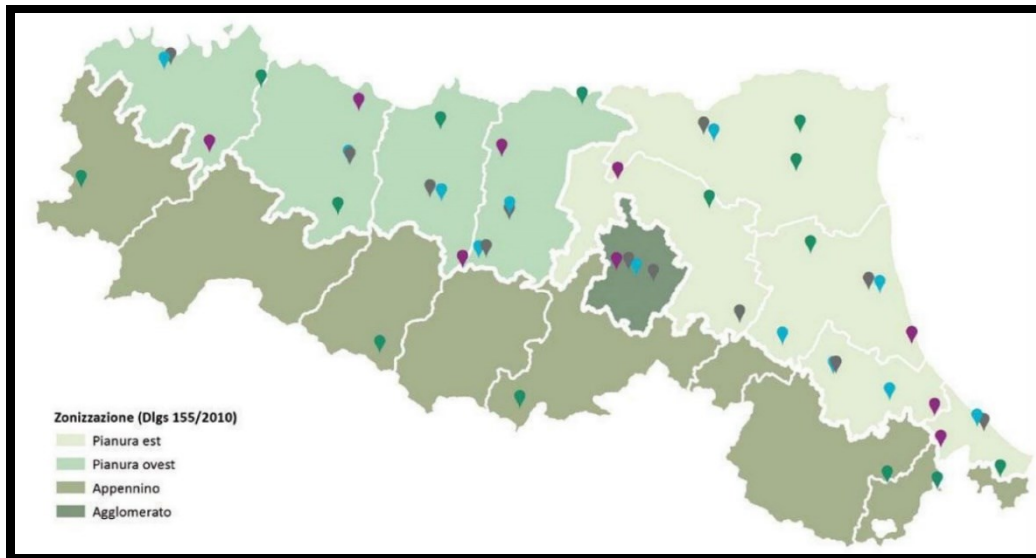


Figura 14: zonizzazione dell'Emilia Romagna. Fonte: ARPA Emilia Romagna.

Le tipologie di stazioni di monitoraggio più comunemente utilizzate dalle regioni sono solitamente quattro:

- Stazione di traffico urbano o UT (Urban Traffic): posizionata a bordo strada, dove il livello di inquinamento è influenzato prevalentemente da emissioni dovute al traffico veicolare. È posta in aree urbane, quindi prevalentemente edificate.
- Stazione di fondo urbano o UB (Urban Background): posizionata dove il livello di inquinamento non è influenzato da una fonte in particolare, ma dal contributo integrato di tutte. È posta in aree urbane, quindi prevalentemente edificate.
- Stazione di fondo suburbano o SB (Suburban Background): posizionata dove il livello di inquinamento non è influenzato da una fonte in particolare, ma dal contributo integrato di tutte. È posta in aree suburbane, solo parzialmente edificate.
- Stazione di fondo rurale o RB (Rural Background): posizionata dove il livello di inquinamento non è influenzato da una fonte in particolare, ma dal contributo integrato di tutte. È posta in aree rurali, quindi in aree distanti dalle fonti di emissione.

Le stazioni che sono state prese in considerazione in questo studio sono state le centraline UB o, nel caso in cui non fossero disponibili nelle provincie delle regioni in questione, quelle SB, perché più rappresentative della concentrazione generale (e non di picco o di minimo) di PM_{10} nell'aria delle provincie considerate. Per l'ozono, sono stati utilizzati i dati rilevati dalle stesse stazioni utilizzate per il PM_{10} .

I dati forniti dall'ARPA Lombardia sono in formato CSV, quindi la prima operazione che si è resa necessaria è stata quella di suddividere il contenuto del file in colonne Excel. Invece i dati forniti dall'ARPA Emilia Romagna sono già riportati in colonne.

Per quanto riguarda la gli inquinanti che sono stati presi in considerazione nell'analisi, la scelta è stata basata sul fatto che essi sono presenti nelle aree di interesse in elevate concentrazioni (spesso al di sopra dei limiti normativi) ed inoltre essi risultano particolarmente pericolosi per la salute della popolazione e possono avere effetti e conseguenze cliniche che potrebbero aver incrementato la mortalità e/o la trasmissione del SARS-CoV-2. In particolare, come detto in precedenza, è stato scelto il PM_{10} piuttosto che il $PM_{2,5}$ perché quest'ultimo è una frazione sul totale del PM_{10} (solitamente compresa tra il 50% ed il 70%).

Le serie di dati prese in considerazione sono relative agli ultimi quattro anni (2016-2020), poiché gli effetti degli inquinanti atmosferici sulla salute umana non dipendono solo dalle loro concentrazioni, ma anche dalla cronicità dell'esposizione. Per quanto riguarda il PM_{10} i valori disponibili sul sito dell'ARPA sono dati come media giornaliera. I dati che sono stati ricavati, elaborati e presi in considerazione nell'analisi sono il numero di superamenti rispetto al limite di legge giornaliero di $50\mu g/m^3$ e al valore di $20\mu g/m^3$ (nuova proposta dell'OMS come concentrazione limite annuale) nei quattro anni considerati e la media complessiva della concentrazione nel periodo considerato. Invece di utilizzare una concentrazione media del periodo considerato per le concentrazioni mancanti (non rappresentativa in questo studio in quanto sono stati considerati anche i superamenti rispetto ai suddetti valori, e quindi valori di picco registrati dalle centraline), si è preferito escludere dall'analisi tutti i giorni con dati non registrati dalle centraline. Per l'ozono invece i valori sono resi disponibili come media oraria, ed il valore ricavato ed utilizzato in questo studio è il numero di superamenti orari rispetto al valore orario di $180\mu g/m^3$ (soglia di informazione) che si sono verificati nel periodo 2016-2020. Analogamente a quanto fatto per il PM_{10} , i dati orari mancanti nelle serie riportate dall'ARPA non sono stati considerati nell'analisi.

La scelta di considerare i superamenti dei valori precedentemente elencati, sia per il PM_{10} che per l'ozono, è dovuto al fatto che i superamenti rispetto a tali valori sono molto frequenti nella Pianura Padana e non solo. Infatti la percentuale di popolazione dell'Unione Europea esposta a livelli superiori ai limiti normativi è il 13-19% per il PM_{10} e il 12-29% per l'ozono (European Environmental Agency, 2019).

Stazione	Data inizi	Data fine	Strumento	PM10	Unità di misura	Media 2016-2020	N. valori > 20 µg/m³	N. valori > 50 µg/m³
GIARDINI MARGHERITA	01/02/2017	02/02/2017	FAI SWAM 5A	238	µg/m³	22.854	596	81
GIARDINI MARGHERITA	31/01/2017	01/02/2017	FAI SWAM 5A	188	µg/m³			
GIARDINI MARGHERITA	02/02/2017	03/02/2017	FAI SWAM 5A	174	µg/m³			
GIARDINI MARGHERITA	18/11/2016	19/11/2016	FAI SWAM 5A	105	µg/m³			
GIARDINI MARGHERITA	30/01/2016	31/01/2016	FAI SWAM 5A	92	µg/m³			
GIARDINI MARGHERITA	30/01/2017	31/01/2017	FAI SWAM 5A	89	µg/m³			
GIARDINI MARGHERITA	27/01/2018	28/01/2018	FAI SWAM 5A	87	µg/m³			
GIARDINI MARGHERITA	03/02/2017	04/02/2017	FAI SWAM 5A	84	µg/m³			
GIARDINI MARGHERITA	26/01/2018	27/01/2018	FAI SWAM 5A	84	µg/m³			
GIARDINI MARGHERITA	04/11/2017	05/11/2017	FAI SWAM 5A	81	µg/m³			
GIARDINI MARGHERITA	08/01/2019	09/01/2019	FAI SWAM 5A	77	µg/m³			
GIARDINI MARGHERITA	29/01/2017	30/01/2017	FAI SWAM 5A	76	µg/m³			
GIARDINI MARGHERITA	12/02/2017	13/02/2017	FAI SWAM 5A	76	µg/m³			
GIARDINI MARGHERITA	26/12/2017	27/12/2017	FAI SWAM 5A	75	µg/m³			
GIARDINI MARGHERITA	03/01/2017	04/01/2017	FAI SWAM 5A	74	µg/m³			
GIARDINI MARGHERITA	17/12/2019	18/12/2019	FAI SWAM 5A	74	µg/m³			
GIARDINI MARGHERITA	14/10/2019	15/10/2019	FAI SWAM 5A	73	µg/m³			
GIARDINI MARGHERITA	21/10/2017	22/10/2017	FAI SWAM 5A	71	µg/m³			
GIARDINI MARGHERITA	23/11/2017	24/11/2017	FAI SWAM 5A	71	µg/m³			
GIARDINI MARGHERITA	22/01/2016	23/01/2016	FAI SWAM 5A	70	µg/m³			
GIARDINI MARGHERITA	08/12/2017	09/12/2017	FAI SWAM 5A	70	µg/m³			
GIARDINI MARGHERITA	09/12/2019	10/12/2019	FAI SWAM 5A	70	µg/m³			
GIARDINI MARGHERITA	21/02/2019	22/02/2019	FAI SWAM 5A	67	µg/m³			
GIARDINI MARGHERITA	17/02/2018	18/02/2018	FAI SWAM 5A	66	µg/m³			
GIARDINI MARGHERITA	27/01/2019	28/01/2019	FAI SWAM 5A	66	µg/m³			

Figura 15: esempio di organizzazione di un foglio Excel per il PM_{10} per la provincia di Bologna, periodo 2016-2020. Nella prima colonna è riportata la posizione della stazione, nella seconda e nella terza la data del rilevamento giornaliero, nella quarta la tipologia di strumento utilizzato per il rilevamento, nella quinta la concentrazione misurata e nella sesta l'unità di misura del dato. Nelle colonne M, N e O sono riportati rispettivamente la media di tutti i valori dei quattro anni, il numero dei valori maggiori di $20\mu\text{g}/\text{m}^3$ e maggiori di $50\mu\text{g}/\text{m}^3$. In questo foglio le concentrazioni rilevate sono state ordinate dalla più elevata alla più bassa.

2.2 Dati di contagio e di mortalità del COVID-19

Le altre categorie di dati che si sono rese necessario per questo studio sono quelle relative al numero dei casi positivi al SARS-CoV-2 e al numero dei decessi causati dal virus. Per quanto riguarda la prima categoria, il numero dei positivi al nuovo Coronavirus è rilevato mediante test del tampone naso-faringeo, che attualmente costituisce l'unico modo sufficientemente veloce ed affidabile per capire se un soggetto ha contratto l'infezione attiva da SARS-CoV-2. Esso consiste nel prelevare un campione di mucosa presente nei primi tratti delle vie respiratorie attraverso un bastoncino cotonato, evitando di entrare in

contatto con denti e lingua. Una volta prelevato il campione, esso viene accuratamente sigillato ed inviato ad un laboratorio di analisi che stabilirà se il soggetto ha contratto l'infezione da Coronavirus o meno in poche ore. I dati, relativi a tutti i casi di COVID-19 diagnosticati microbiologicamente (tampone naso-faringeo) provenienti dai laboratori di riferimento regionali, vengono raccolti dalle regioni/province autonome ed inviati giornalmente al Dipartimento della Protezione Civile ed al Ministero della Salute. Solitamente i dati dei positivi rilevati con il test del tampone naso-faringeo sono relativi ai 5-7 giorni precedenti (periodo di incubazione medio) alla rilevazione del virus nel soggetto.

Questo test non va confuso con i test sierologici che invece misurano la presenza di anticorpi al SARS-CoV-2 nel sangue, e quindi viene effettuato solo nei soggetti che hanno superato la fase dell'infezione attiva.

Al Dipartimento della Protezione Civile è stato affidato il coordinamento degli interventi necessari a fronteggiare l'emergenza Coronavirus sul territorio nazionale. Inoltre alla Protezione Civile è stata affidata anche la stipulazione dei contratti relativi all'acquisto delle attrezzature mediche necessarie a fronteggiare la pandemia (mascherine chirurgiche, ventilatori polmonari...).

Detto ciò, alcuni dei dati utilizzati nell'analisi sono stati scaricati attraverso il database realizzato dal Dipartimento della Protezione Civile, sulla base dei dati forniti dal Ministero della Salute, e che riporta i seguenti numeri:

- Attualmente positivi, a livello nazionale e regionale;
- Guariti, a livello nazionale;
- Deceduti, a livello nazionale e regionale (dato provvisorio, in attesa di verifica dall'ISS);
- Totale positivi, a livello nazionale, regionale e provinciale;
- Incremento giornaliero di tutti i dati sopraindicati.

Da questo database è stato possibile scaricare il numero totale dei positivi per provincia, disponibile a partire dal 24 febbraio 2020. Poiché i rilevamenti dei primi positivi da parte della protezione civile sono iniziati solo a fine febbraio, quando già erano presenti casi di infezione, i primi dati riportati nel database sono sicuramente meno rappresentativi di quelli dei mesi successivi. Inoltre anche nel mese di marzo il rapporto nuovi casi/tamponi risulta essere molto elevato rispetto ai mesi successivi, fatto che conferma anche che nei

primi periodi della pandemia il numero di tamponi non è stato sufficientemente elevato, come si può vedere dalle seguenti figure.

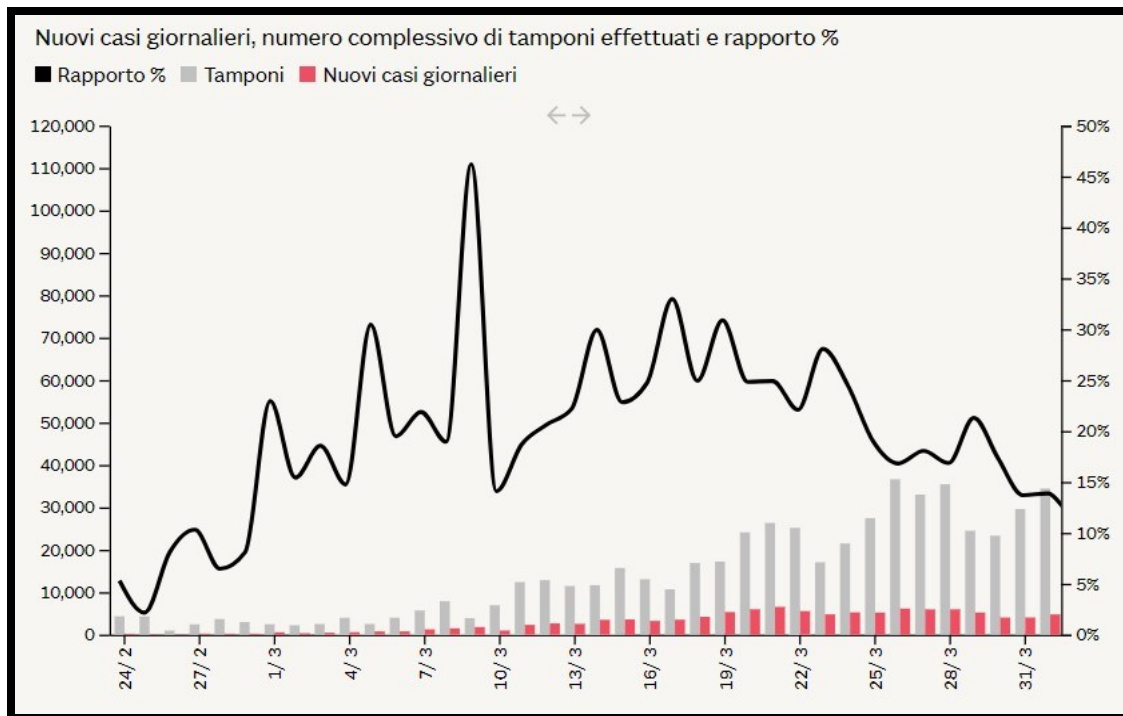


Figura 16: andamento del rapporto nuovi casi/tamponi, del numero di tamponi e dei nuovi casi nel periodo fine febbraio-fine marzo 2020. Si nota chiaramente che il rapporto nuovi casi/tamponi eseguiti risulta particolarmente elevato, con picchi di quasi il 50%. Fonte: Lab24 de “Il Sole 24 Ore”.

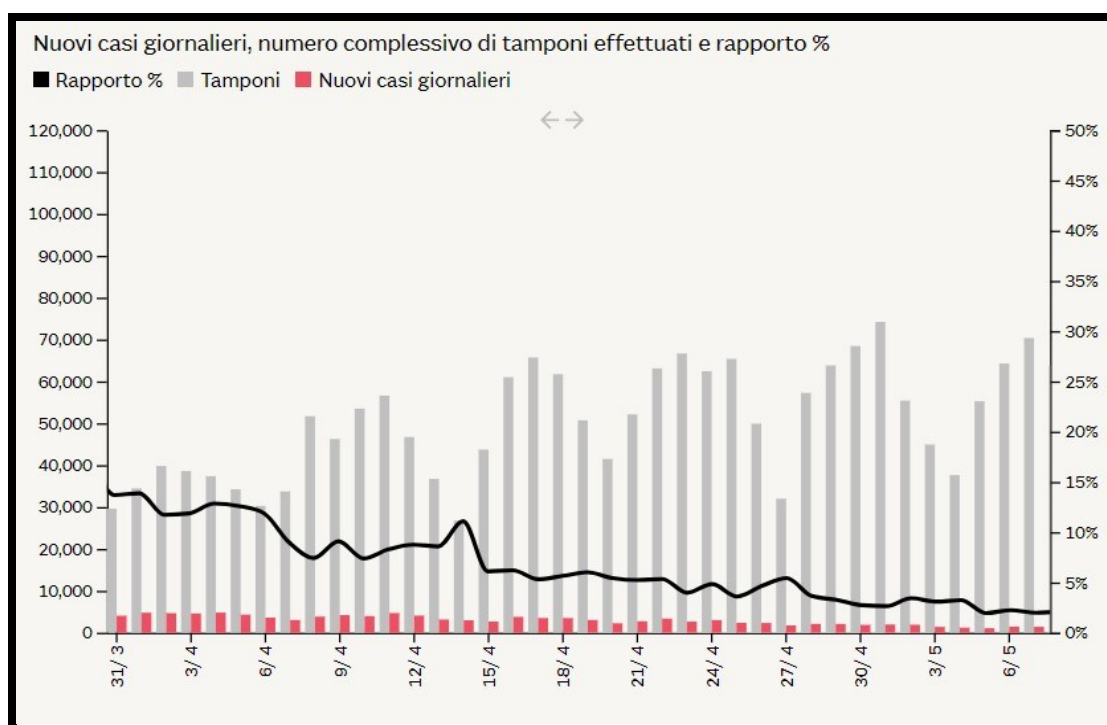


Figura 17: andamento del rapporto nuovi casi/tamponi, del numero di tamponi e dei nuovi casi nel periodo fine marzo-inizio maggio 2020. Si nota subito che il rapporto nuovi casi/tamponi eseguiti si è drasticamente ridotto nei mesi successivi rispetto ai mesi di febbraio e marzo. Fonte: Lab24 de "Il Sole 24 Ore".

Questo fatto potrebbe anche spiegare (perlomeno in parte, in quanto non si escludono anche altri fattori che possano aver contribuito come, ad esempio, elevati livelli di inquinamento, elevata età media...) l'elevata insolita mortalità del SARS-CoV-2 che è stata registrata soprattutto nei primi mesi della pandemia in Lombardia. Infatti, come si può vedere nella figura seguente, il rapporto nuovi casi/tamponi eseguiti nel mese di marzo in Lombardia risulta estremamente elevato, con picchi dell'80%.

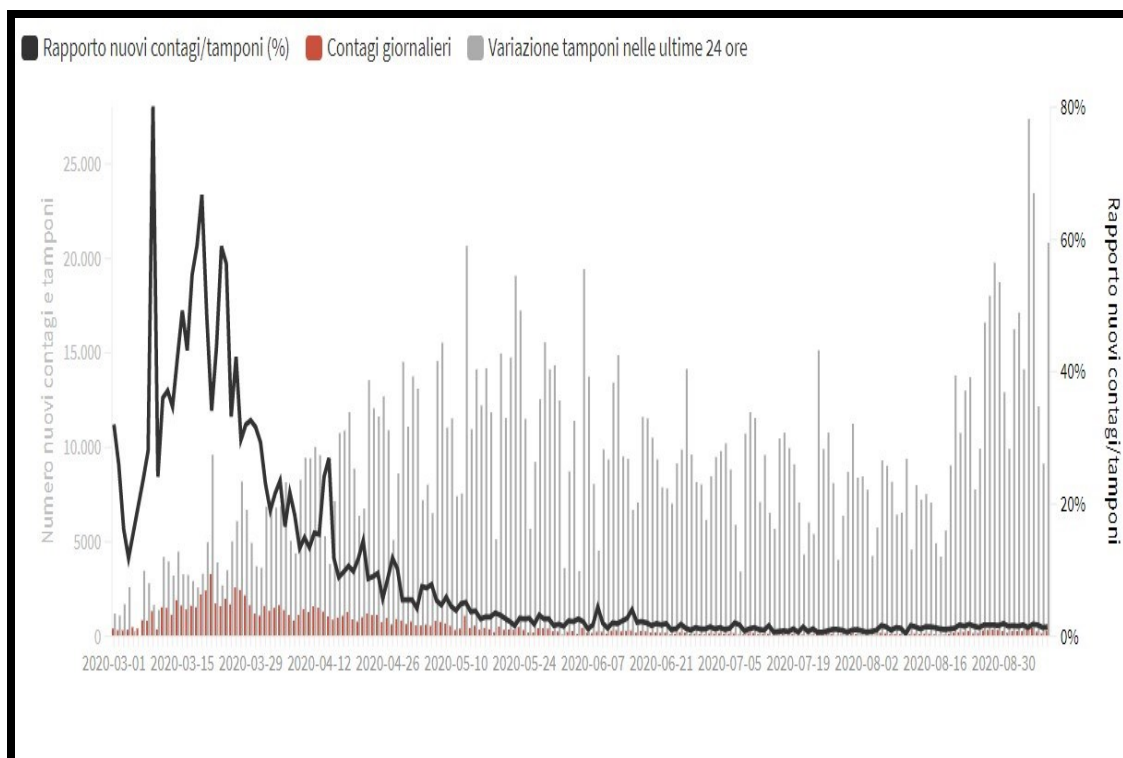


Figura 18: andamento del rapporto nuovi casi/tamponi, del numero di tamponi e dei nuovi casi nel periodo fine febbraio-fine agosto 2020 in Lombardia. Si vede chiaramente che nel periodo di febbraio e marzo il numero di tamponi è sensibilmente più basso rispetto ai mesi successivi e quindi il rapporto nuovi contagi/tamponi risulta molto elevato in questo periodo. Fonte: GEDI Visual.

Per quanto riguarda il numero dei decessi per COVID-19, fino al 31 marzo ed al 30 aprile 2020 nelle province della Lombardia e dell'Emilia Romagna, è stato estrapolato dal primo e dal secondo rapporto ISTAT-ISS, pubblicati rispettivamente nei mesi di maggio e giugno. Il numero di decessi indicato nei rapporti è stato esplicitamente confermato dall'ISS. Però nell'analisi, come dato di mortalità del SARS-CoV-2, si è deciso di non utilizzare soltanto il numero di decessi, ma il numero dei decessi diviso il numero di casi totali che si sono verificati in ogni provincia delle regioni indagate. Lo scopo di questa scelta è quello di avere a disposizione un dato più oggettivo e con minore variazione, perché indipendente dal numero dei casi di COVID-19 che si sono verificati nelle diverse province. Il parametro utilizzato è stato ricavato dalla formula seguente.

$$X_{i,j,k} = D_{i,j,k} / C_{i,j,k}, \text{ dove:}$$

- $X_{i,j,k}$ è il parametro di mortalità per COVID-19 per la i -esima provincia della j -esima regione, fino al k -esimo giorno.

- $D_{i,j,k}$ è il numero di decessi dovuti al COVID-19 per la i -esima provincia della regione j -esima, fino al k -esimo giorno.
- $C_{i,j,k}$ è il numero di casi totali di COVID-19 nella i -esima provincia della j -esima regione, fino al k -esimo giorno.

3. Risultati

In questo capitolo verranno presentati tutti i risultati numerici e grafici ottenuti con l'elaborazione statistica dei dati di inquinamento atmosferico e quelli relativi al COVID-19, descritti nel dettaglio nel capitolo precedente. Per elaborare le serie di dati considerate, si è fatto uso della regressione lineare, che consiste in un metodo di stima del valore atteso condizionato di una variabile dipendente (Y), dati i valori di altre variabili indipendenti ($X_1, X_2, \dots, X_K$). In particolare è stato utilizzato il modello di regressione lineare semplice:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

che identifica una retta nota come retta di regressione lineare semplice, dove:

- y_i è la variabile dipendente;
- x_i è la variabile indipendente;
- β_0 è l'intercetta della retta di regressione semplice;
- β_1 è il coefficiente angolare della retta di regressione semplice;
- ε_i è l'errore statistico.

Ogni retta di regressione individuata in quest'analisi è contraddistinta dal coefficiente di determinazione (r^2). Quest'indice statistico è una proporzione tra la variabilità dei dati e la correttezza del modello statistico utilizzato e misura la frazione della varianza della variabile dipendente espressa dalla regressione. Questo coefficiente varia tra 0 ed 1: quando è 0 il modello utilizzato non spiega per nulla i dati; quando è 1 il modello spiega perfettamente i dati. Nelle regressioni lineari semplici r^2 è semplicemente il quadrato del coefficiente di correlazione (anche detto coefficiente di Pearson).

Le variabili considerate che sono state sottoposte a questo tipo di correlazione statistica, per tutte le provincie di Lombardia ed Emilia Romagna, sono riportate nella seguente tabella.

Tabella 1: variabili utilizzate nel modello di regressione lineare per la valutazione della correlazione statistica.

Variabile indipendente (x_i)	Variabile dipendente (y_i)
Media PM ₁₀ 2016-2020	Decessi/casi totali al 31/03/2020
Media PM ₁₀ 2016-2020	Decessi/casi totali al 30/04/2020

Media PM ₁₀ 2016-2020	Casi totali al 28/05/2020
N. valori di PM ₁₀ > 20µg/m ³ 2016-2020	Decessi/casi totali al 31/03/2020
N. valori di PM ₁₀ > 20µg/m ³ 2016-2020	Decessi/casi totali al 30/04/2020
N. valori di PM ₁₀ > 20µg/m ³ 2016-2020	Casi totali al 28/05/2020
N. valori di PM ₁₀ > 50µg/m ³ 2016-2020	Decessi/casi totali al 31/03/2020
N. valori di PM ₁₀ > 50µg/m ³ 2016-2020	Decessi/casi totali al 30/04/2020
N. valori di PM ₁₀ > 50µg/m ³ 2016-2020	Casi totali al 28/05/2020
N. valori di O ₃ > 180µg/m ³ 2016-2020	Decessi/casi totali al 31/03/2020
N. valori di O ₃ > 180µg/m ³ 2016-2020	Decessi/casi totali al 30/04/2020
N. valori di O ₃ > 180µg/m ³ 2016-2020	Casi totali al 28/05/2020

Una volta ottenuti i grafici di dispersione, costruiti riportando sull'asse delle x i valori assunti dalla variabile indipendente e sull'asse delle y quelli della variabile dipendente, con la relativa retta di regressione lineare semplice, è stata valutata la correlazione statistica tra le due variabili utilizzando il coefficiente di correlazione di Pearson “r” (anche detto coefficiente di correlazione lineare), un indice statistico che esprime un'eventuale relazione di linearità tra esse. Questo coefficiente può assumere tutti i valori

compresi tra -1 (perfetta correlazione lineare inversa) ed 1 (perfetta correlazione lineare); si possono distinguere tre casi differenti:

- $r < 0$, se le variabili sono inversamente correlate;
- $r = 0$, se le variabili non sono tra loro correlate;
- $r > 0$, se le variabili sono direttamente correlate.

Per valutare la significatività statistica delle serie di dati utilizzate, si è utilizzato il parametro p-value che stabilisce se la significatività tra il risultato osservato e quello ipotizzato è dovuta alla casualità introdotta dal campionamento, oppure se tale differenza è statisticamente significativa, cioè difficilmente spiegabile mediante la casualità dovuta al campionamento. Per la conduzione di questa analisi si è fissato il valore soglia di significatività “ α ” pari al 5% (0,05), convenzionalmente assunto in questo tipo di verifica. Calcolato il p-value per le serie di dati in esame, i casi che si possono presentare sono due:

- se $p\text{-value} > \alpha$ l’evidenza empirica non è sufficientemente contraria all’ipotesi nulla che quindi non può essere rifiutata;
- se $p\text{-value} \leq \alpha$ l’evidenza empirica è fortemente contraria all’ipotesi nulla che quindi va rifiutata. In tal caso si dice che i dati osservati sono statisticamente significativi.

3.1 Risultati per la regione Lombardia

In questo paragrafo verranno presentati tutti i risultati delle correlazioni, precedentemente elencate in tabella 1, solamente per la regione Lombardia. I risultati sono stati ottenuti tramite regressione lineare dei dati relativi al PM_{10} (medie e superamenti dei valori descritti nel capitolo precedente) e quelli relativi al SARS-CoV-2 (casi totali e decessi/casi totali) delle 12 provincie della regione. Ogni retta di regressione lineare è quindi stata ottenuta interpolando i 12 punti nel grafico, ognuno corrispondente ad una provincia della regione Lombardia. I primi risultati che verranno presentati sono quindi relativi alla media per il 2016-2020 per le varie provincie.

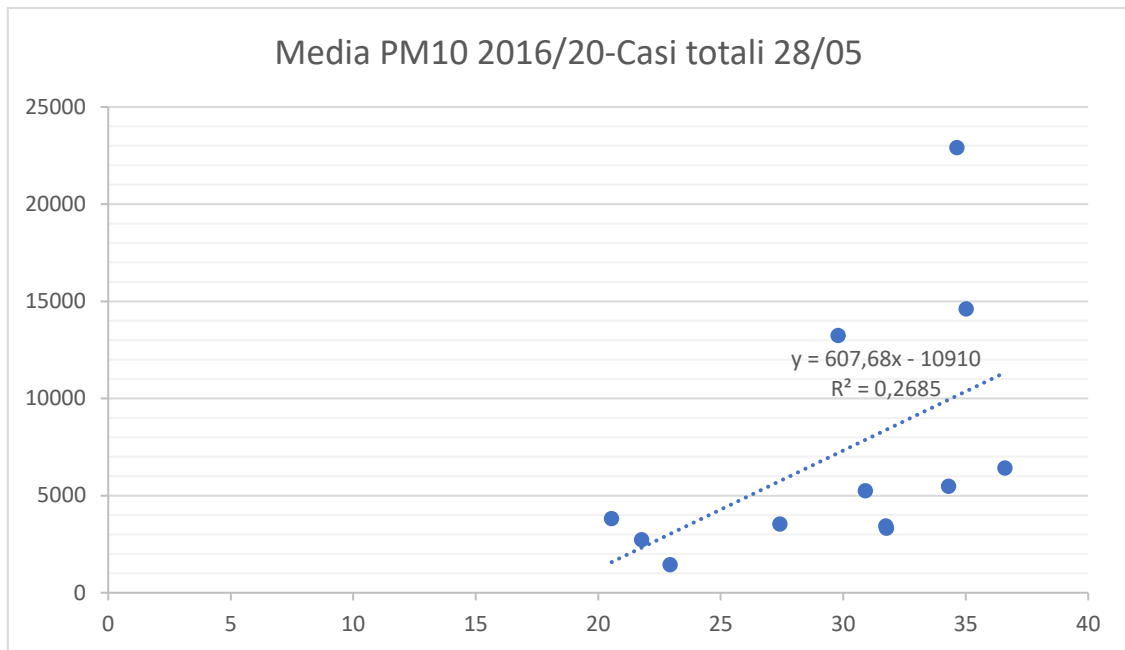


Grafico 1: retta di regressione lineare ottenuta per interpolazione dei punti media PM_{10} (asse x) e casi totali al 28/05/2020 (asse y) per le 12 provincie lombarde. In questo primo caso si ha un coefficiente $r^2=0,2685$, un **coefficiente di Pearson** $r=0,5182$ e $p\text{-value}<0,05$.

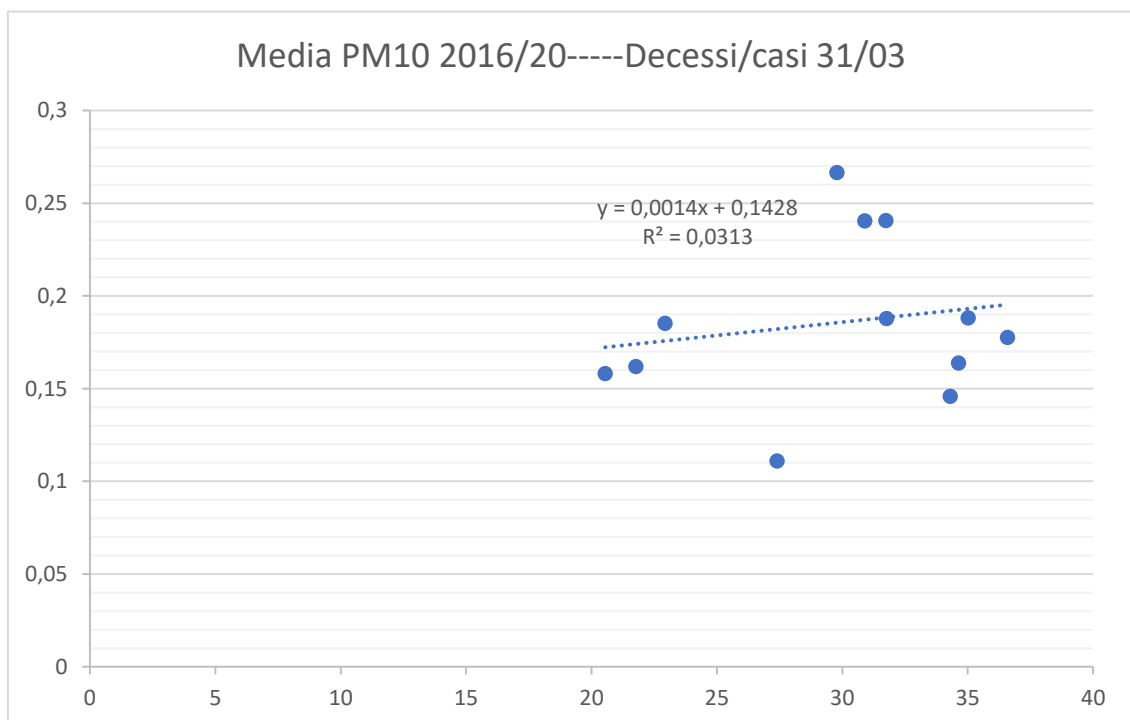


Grafico 2: retta di regressione lineare ottenuta per interpolazione dei punti media PM_{10} (asse x) e decessi/casi totali al 31/03/2020 (asse y) per le 12 provincie lombarde. Il coefficiente r^2 è pari a 0,0313, il **coefficiente di Pearson** è $r=0,1769$ e $p\text{-value}<0,05$.

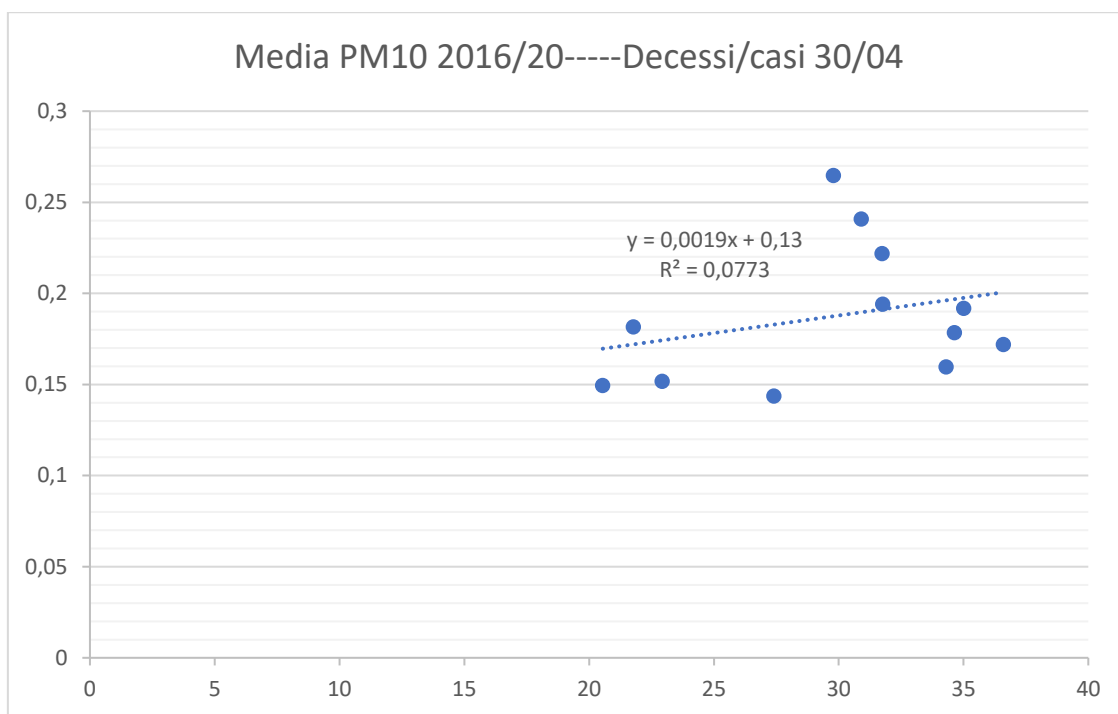


Grafico 3: retta di regressione lineare ottenuta per interpolazione dei punti media PM_{10} (asse x) e decessi/casi totali al 30/04/2020 (asse y) per le 12 provincie lombarde. Si ha un coefficiente $r^2=0,0773$, un **coefficiente di Pearson** $r=0,2780$ e $p\text{-value}<0,05$.

Tutte e tre le analisi risultano statisticamente significative in quanto $p\text{-value}<0,05$. Tuttavia solo la prima retta di regressione lineare è caratterizzata da un coefficiente di Pearson che indica una correlazione tra le variabili degna di nota, ovvero una presunta relazione tra i livelli medi di PM_{10} nei quattro anni precedenti e i casi totali di COVID-19 nelle provincie. Dai successivi grafici 2 e 3 si nota una debole correlazione positiva tra le variabili, in quanto in entrambi i casi si ha un coefficiente di Pearson maggiore di zero, tuttavia non molto elevato.

Per quanto riguarda i risultati ottenuti dall'elaborazione del numero di superamenti per provincia rispetto alla concentrazione di PM_{10} giornaliera di $20\mu\text{g}/\text{m}^3$ che si sono verificati negli anni 2016-2020, i grafici ottenuti sono di seguito riportati.

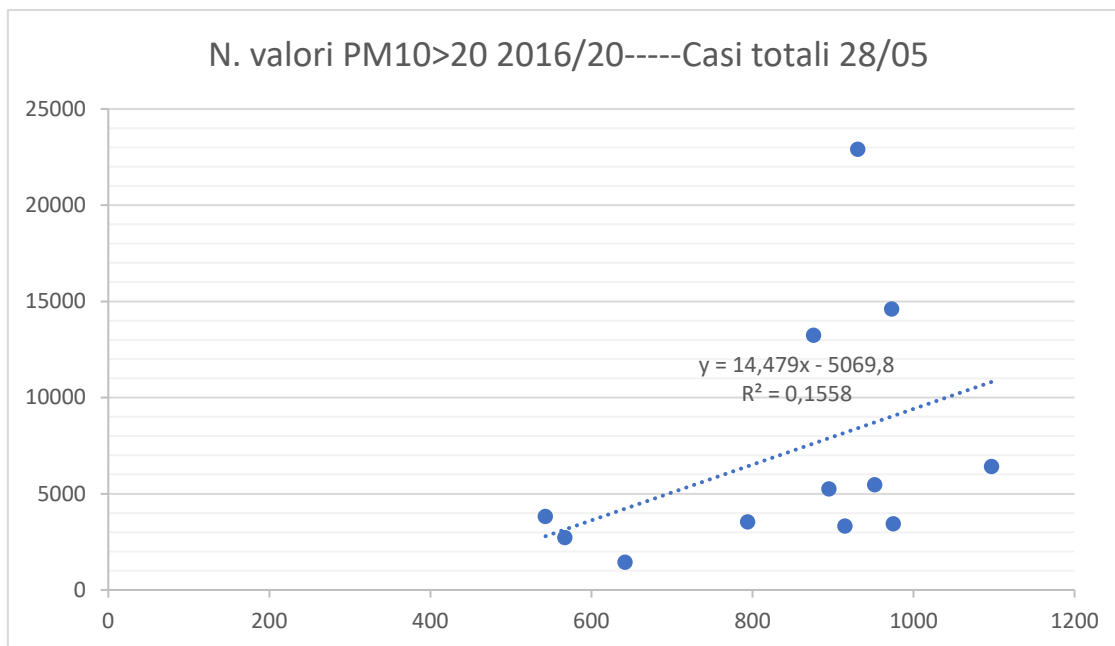


Grafico 4: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $PM_{10} > 20 \mu g/m^3$ nel periodo 2016-2020 (asse x) e casi totali al 28/05/2020 (asse y) per le 12 provincie lombarde. Si ha un coefficiente $r^2=0,1558$, un **coefficiente di Pearson** $r=0,3948$ e $p\text{-value}<0,05$.

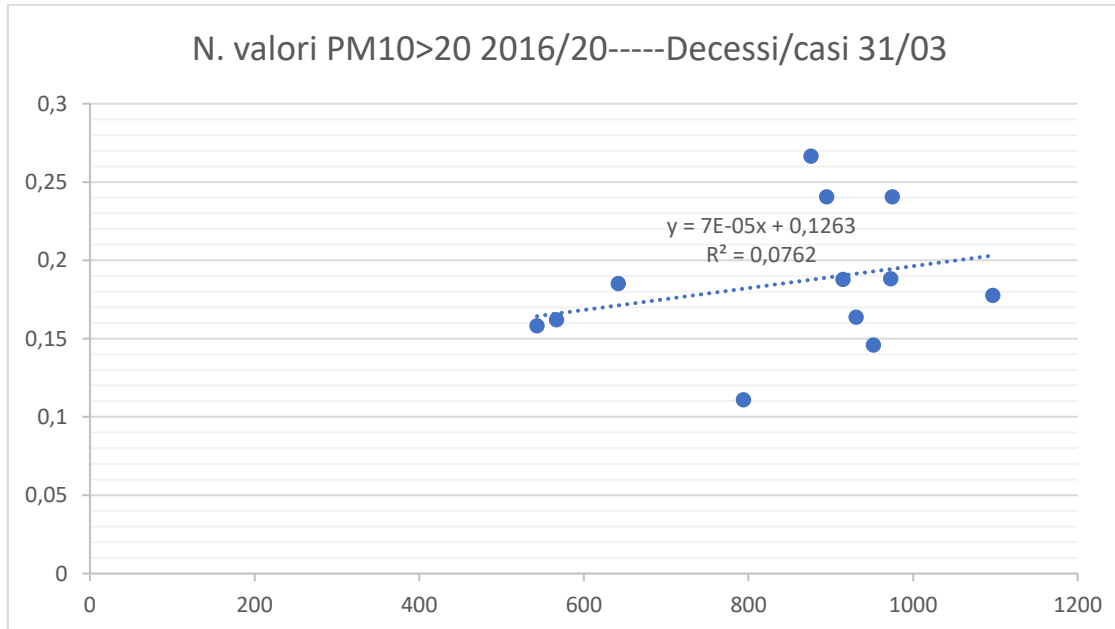


Grafico 5: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $PM_{10} > 20 \mu g/m^3$ nel periodo 2016-2020 (asse x) decessi/casi totali al 31/03/2020 (asse y) per le 12 provincie lombarde. Si ha un coefficiente $r^2=0,0762$, un **coefficiente di Pearson** $r=0,2760$ e $p\text{-value}<0,05$.

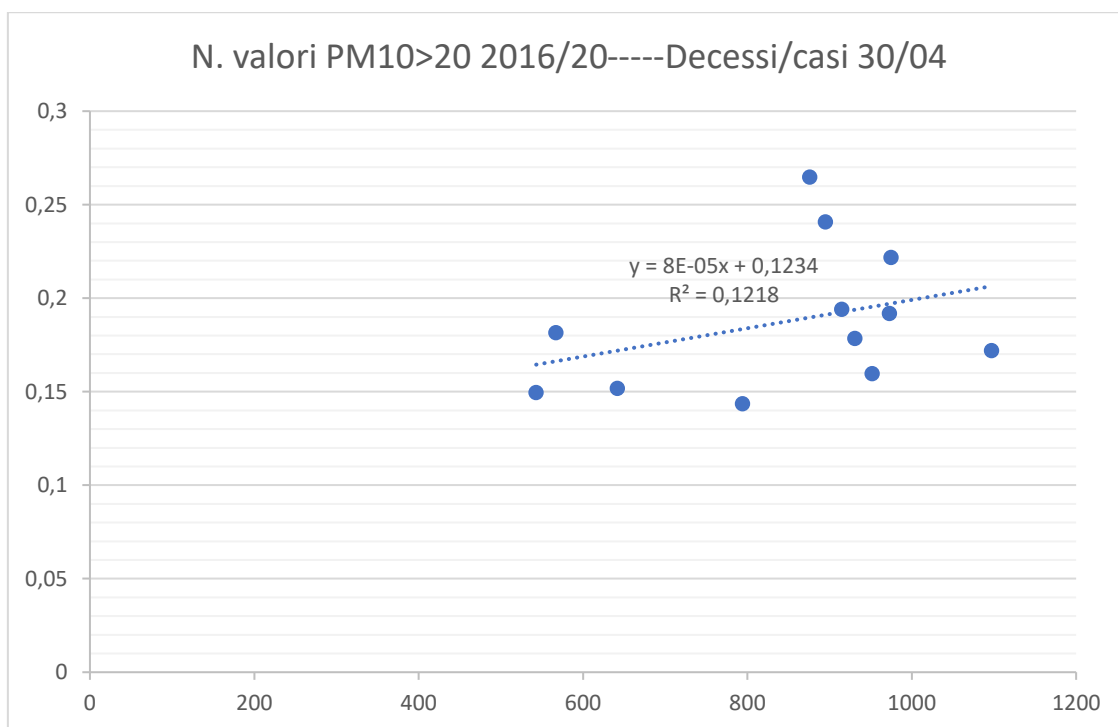


Grafico 6: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $PM_{10} > 20 \mu g/m^3$ nel periodo 2016-2020 (asse x) e decessi/casi totali al 30/04/2020 (asse y) per le 12 provincie lombarde. Si ha un coefficiente $r^2=0,1218$, un **coefficiente di Pearson $r=0,3490$** e $p\text{-value} < 0,05$.

Da questi grafici 4,5 e 6 si nota una diretta proporzionalità tra le due variabili, tuttavia limitata in quanto il coefficiente di Pearson non è particolarmente elevato in nessuno dei tre casi (risulta sempre inferiore a 0,4).

Gli ultimi risultati ricavati per la regione Lombardia, per quanto riguarda il PM_{10} , sono stati ottenuti con l'elaborazione del numero di superamenti giornalieri del limite normativo di $50 \mu g/m^3$ nei quattro anni in questione. I risultati ottenuti sono di seguito riportati.

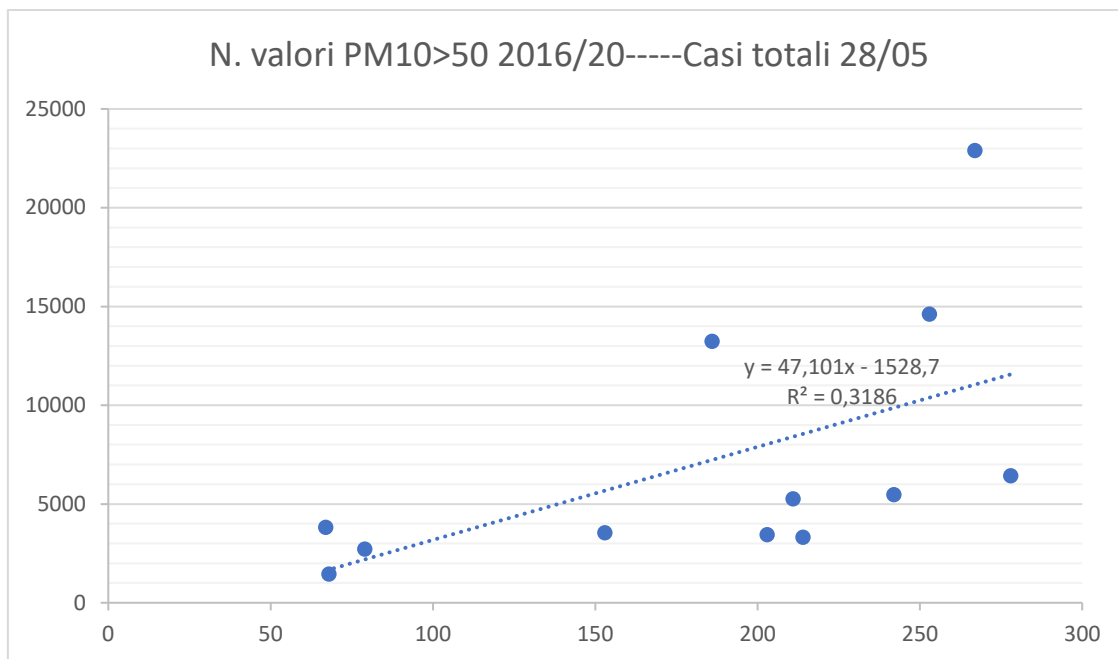


Grafico 7: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $PM_{10} > 50 \mu g/m^3$ nel quadriennio 2016-2020 (asse x) e casi totali al 28/05/2020 (asse y) per le 12 provincie lombarde. Si ha un coefficiente $r^2=0,3186$, un **coefficiente di Pearson** $r=0,5645$ e $p\text{-value}<0,05$.

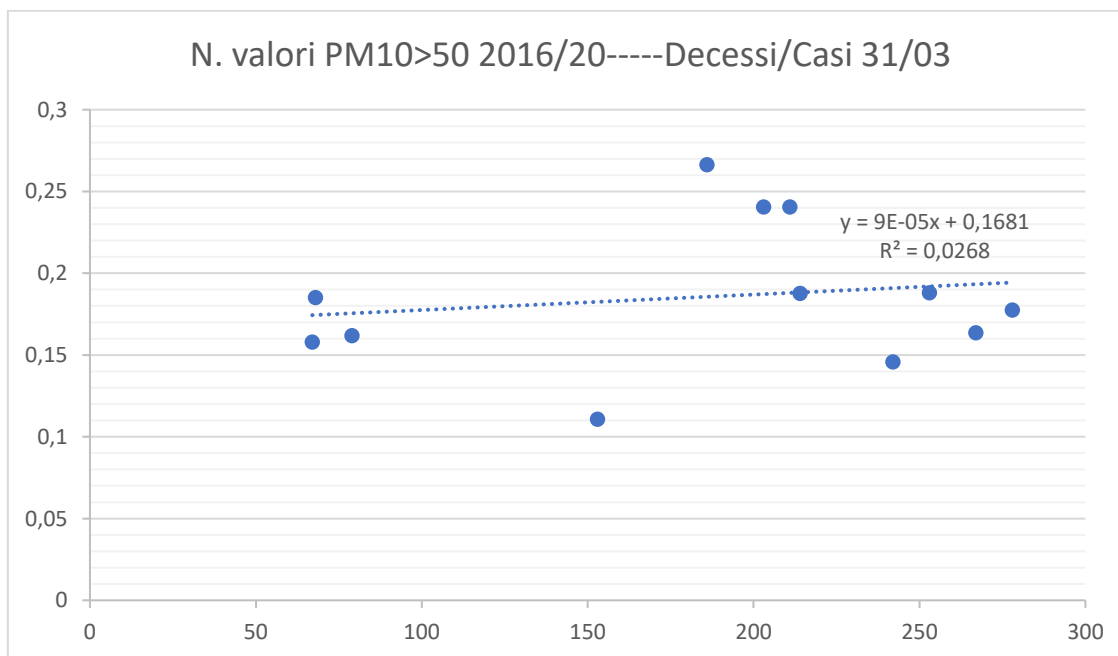
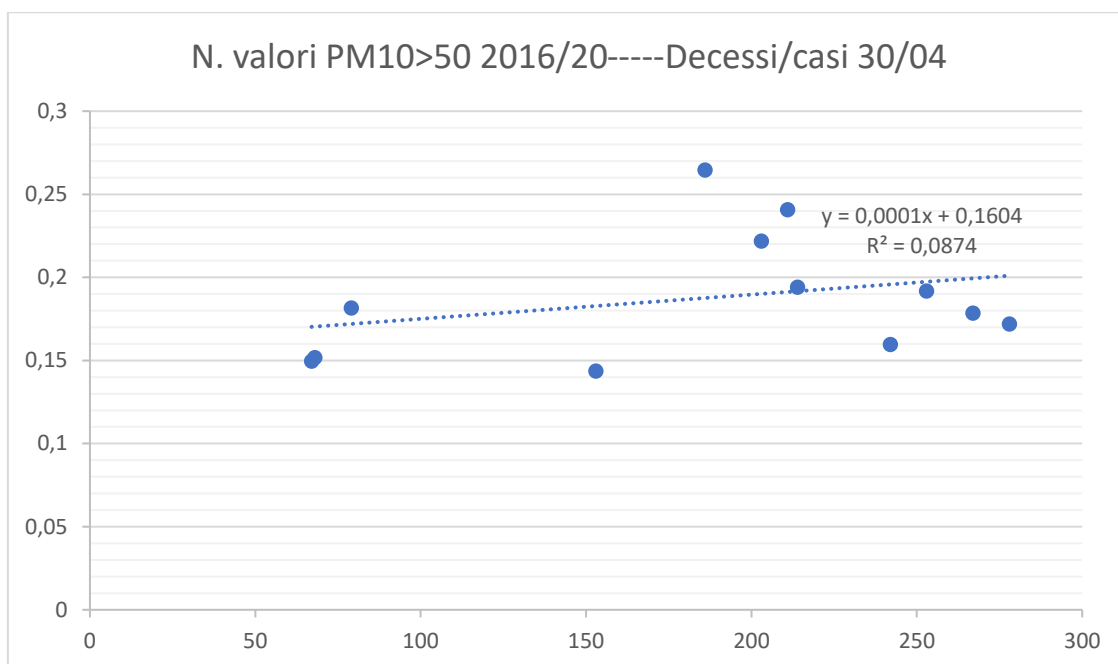


Grafico 8: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $PM_{10} > 50 \mu g/m^3$ nel quadriennio 2016-2020 (asse x) e decessi/casi totali al 31/03/2020 (asse y) per le 12 provincie lombarde. Si ha un coefficiente $r^2=0,0268$, un **coefficiente di Pearson** $r=0,1638$ e $p\text{-value}<0,05$.



*Grafico 9: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $PM_{10} > 50 \mu g/m^3$ nel quadriennio 2016-2020 (asse x) e decessi/casi totali al 30/04/2020 (asse y) per le 12 province lombarde. Si ha un coefficiente $r^2 = 0,0874$, un **coefficiente di Pearson** $r = 0,2957$ e $p\text{-value} < 0,05$.*

In questi ultimi tre grafici soltanto nel grafico 7 la retta di regressione lineare è caratterizzata da un coefficiente di Pearson degno di nota (maggiore di 0,5) e quindi da una diretta proporzionalità tra le due variabili non trascurabile. Invece le rette dei grafici 8 e 9 presentano un coefficiente di Pearson maggiore di zero, tuttavia non molto elevato (inferiore a 0,3 in entrambi i casi) e quindi si ha una debole correlazione lineare tra le variabili considerate.

Gli altri risultati, ottenuti invece considerando il numero di superamenti orari della soglia di informazione ($180 \mu g/m^3$) dell'ozono troposferico nel periodo 2016-2020, sono i seguenti:

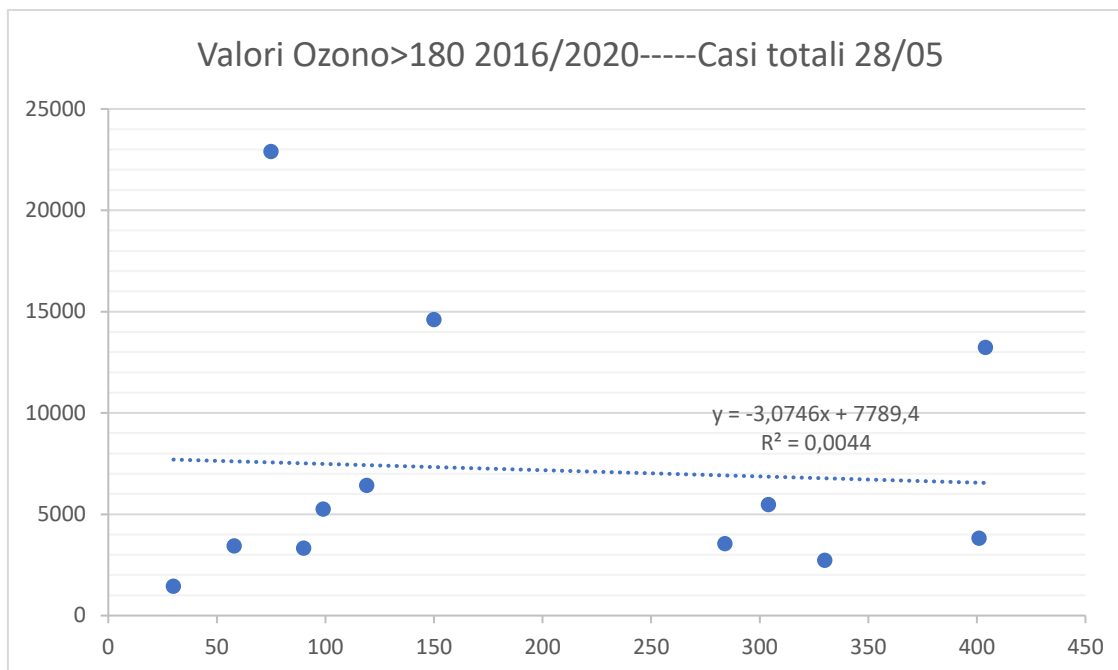


Grafico 10: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $O_3 > 180 \mu\text{g}/\text{m}^3$ nel quadriennio 2016-2020 (asse x) e casi totali al 28/05/2020 (asse y) per le 12 province della Lombardia. Si ha un coefficiente $r^2=0,0044$, un **coefficiente di Pearson $r=-0,0667$** e $p\text{-value}<0,05$.

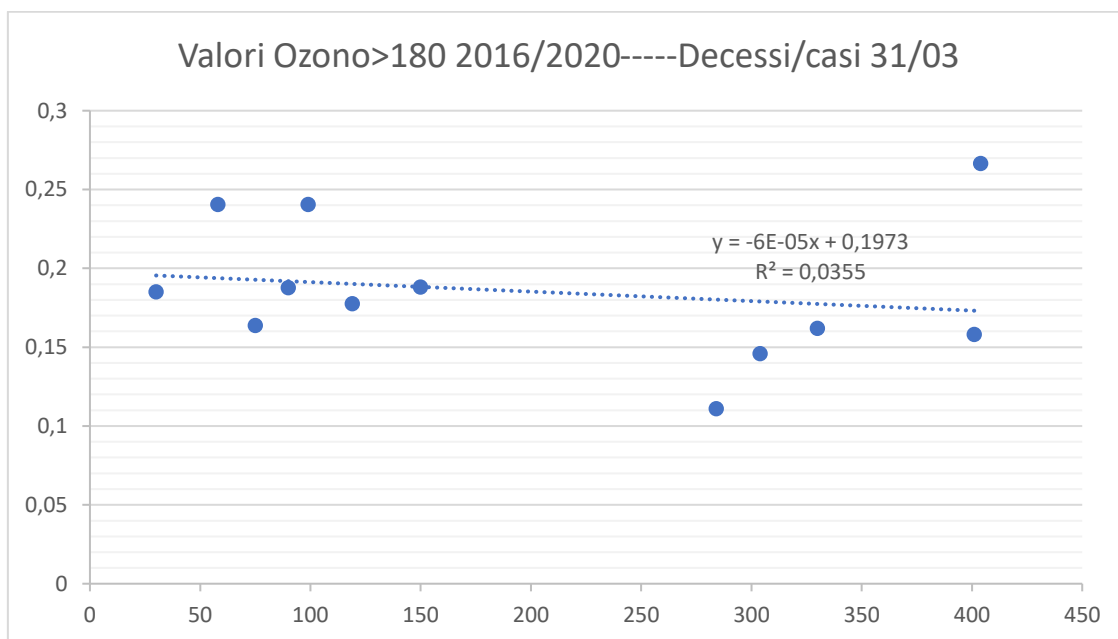


Grafico 11: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $O_3 > 180 \mu\text{g}/\text{m}^3$ nel quadriennio 2016-2020 (asse x) e decessi/casi totali al 31/03/2020 (asse y) per le 12 province della Lombardia. Si ha un coefficiente $r^2=0,0355$, un **coefficiente di Pearson $r=-0,1885$** e $p\text{-value}<0,05$.

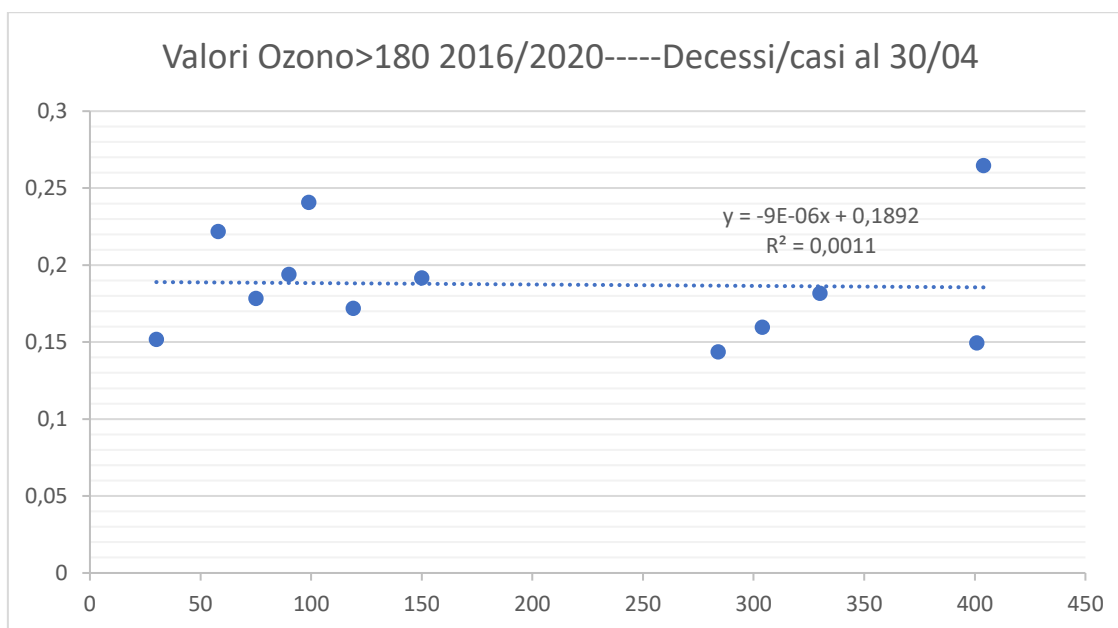


Grafico 12: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $O_3 > 180 \mu\text{g}/\text{m}^3$ nel quadriennio 2016-2020 (asse x) e decessi/casi totali al 30/04/2020 (asse y) per le 12 provincie della Lombardia. Si ha un coefficiente $r^2 = 0,0011$, un coefficiente di Pearson $r = -0,0336$ e $p\text{-value} < 0,05$.

In conclusione, dai risultati ottenuti per la regione Lombardia si può dire che per il PM_{10} emerge una buona correlazione statistica, sia per quanto riguarda la media dei quattro anni passati sia che per il numero di superamenti del valore normativo giornaliero ($50 \mu\text{g}/\text{m}^3$), con il numero di casi che si sono verificati fino al 28 maggio 2020 (e quindi sulla trasmissione del SARS-CoV-2). Invece non è stata riscontrata nessuna particolare correlazione tra i valori di PM_{10} analizzati e la mortalità del virus (decessi/casi totali), a meno di un debole diretta proporzionalità.

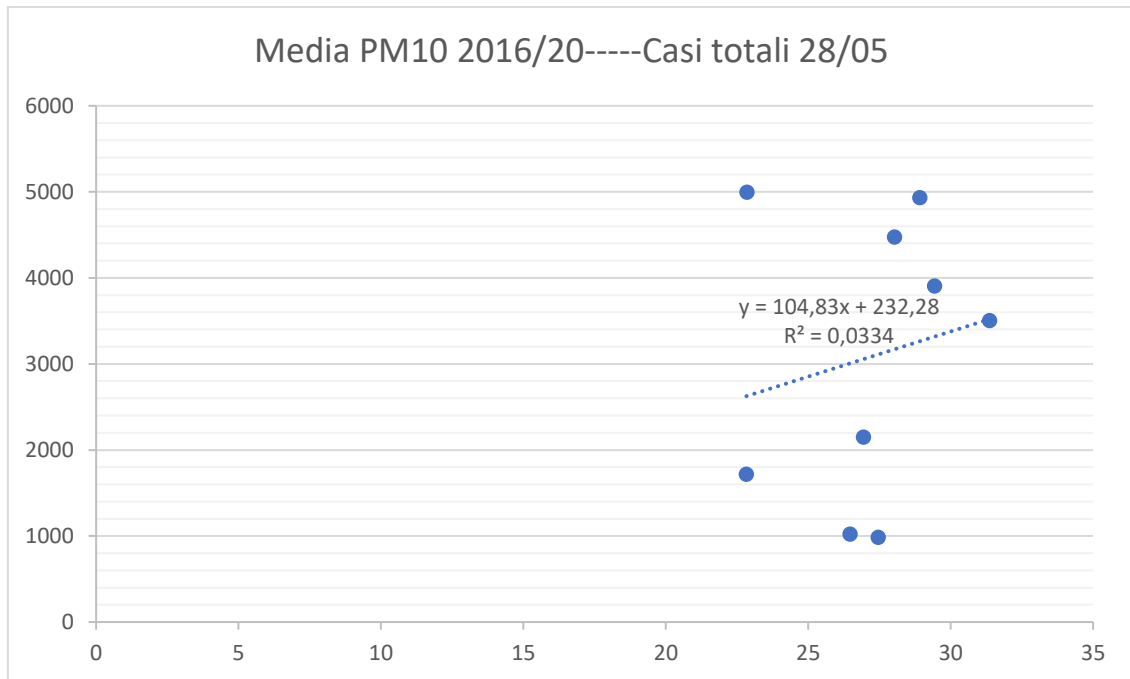
Il discorso invece cambia per l'ozono troposferico per cui non emerge nessun risultato rilevante, in quanto le rette di regressione lineare dei grafici 10, 11 e 12 sono caratterizzate da un coefficiente di Pearson debolmente negativo, e quindi in contraddizione con i risultati ottenuti elaborando i dati relativi al PM_{10} .

3.2 Risultati per la regione Emilia Romagna

In questo paragrafo verranno presentati, in modo analogo a quanto fatto per la regione Lombardia, tutti i risultati ottenuti elaborando le serie di dati delle nove provincie

dell'Emilia Romagna. Ogni retta di regressione è stata quindi ottenuta interpolando 9 punti, ognuno corrispondente ad una provincia.

Anche in questo caso si inizierà quindi esponendo i grafici ricavati elaborando la media di PM_{10} degli ultimi quattro anni.



*Grafico 13: retta di regressione lineare ottenuta per interpolazione dei punti media PM_{10} (asse x) e casi totali al 28/05/2020 (asse y), per le 9 provincie dell'Emilia Romagna. In questa prima analisi si ha un coefficiente $r^2=0,0334$, un **coefficiente di Pearson $r=0,1829$** e $p\text{-value}<0,05$.*

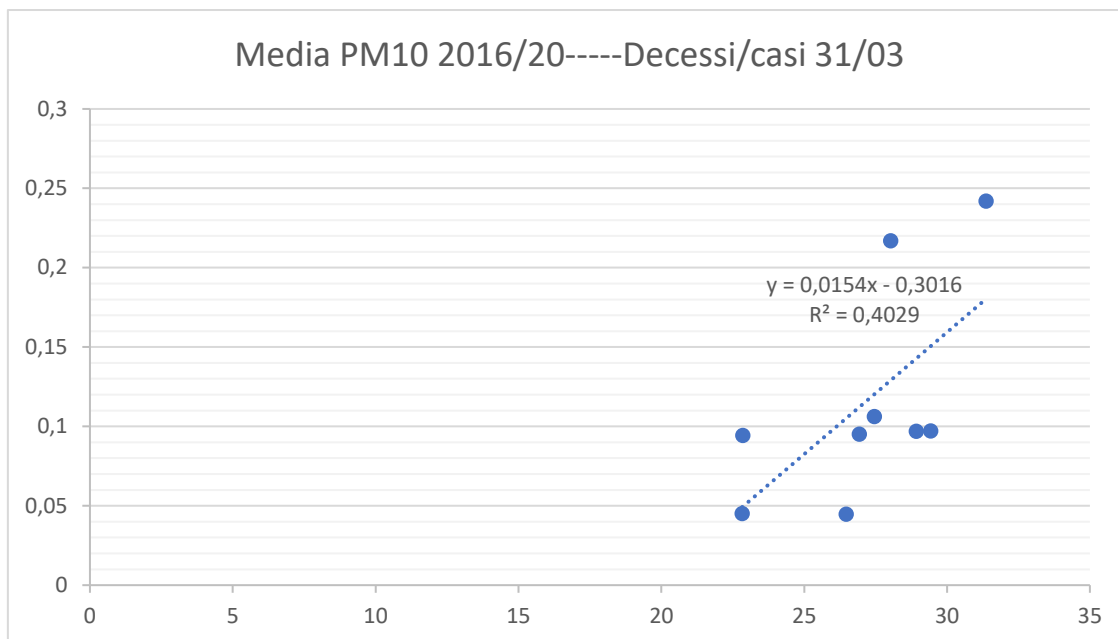


Grafico 14: retta di regressione lineare ottenuta per interpolazione dei punti media PM_{10} (asse x) e decessi/casi totali al 31/03/2020 (asse y), per le 9 province dell'Emilia Romagna. Il coefficiente r^2 è pari a 0,4029, il **coefficiente di Pearson** è $r=0,6347$ e $p\text{-value}<0,05$.

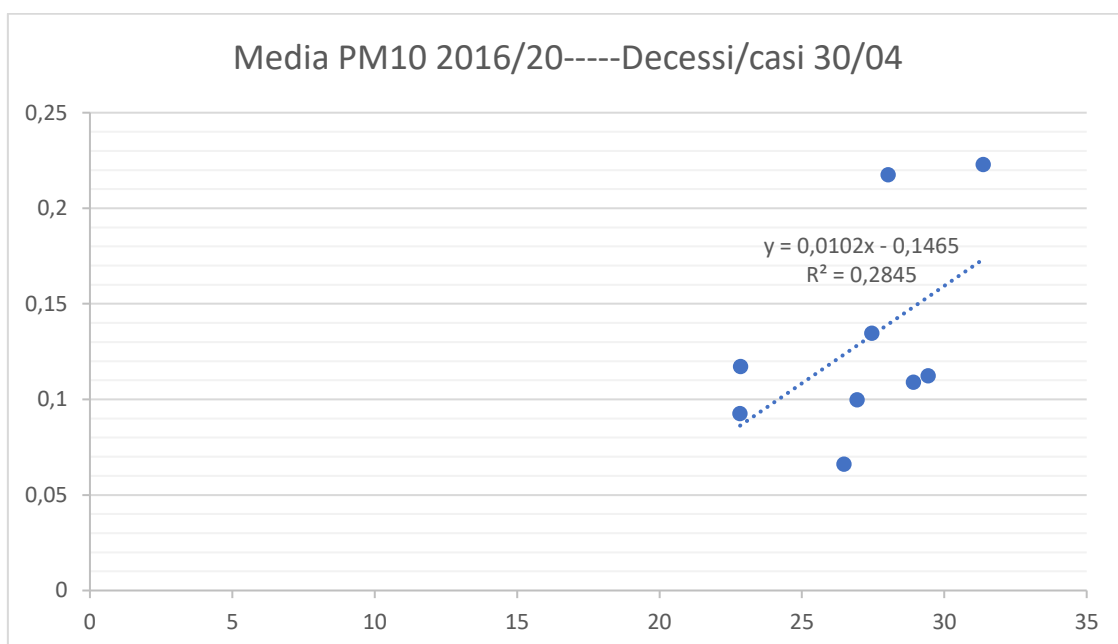
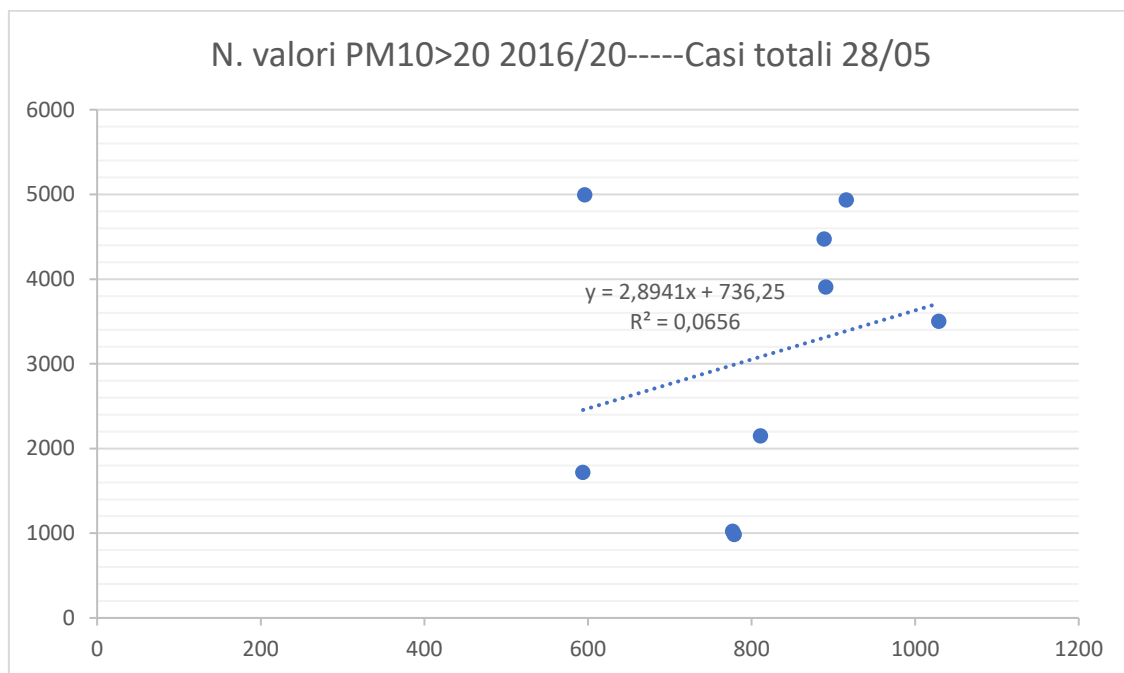


Grafico 15: retta di regressione lineare ottenuta per interpolazione dei punti media PM_{10} (asse x) e decessi/casi totali al 30/04/2020 (asse y), per le 9 province dell'Emilia Romagna. Il coefficiente r^2 è pari a 0,2845, il **coefficiente di Pearson** è $r=0,5334$ e $p\text{-value}<0,05$.

In questi tre grafici è possibile notare una buona correlazione tra la media del PM_{10} nel periodo 2016-2020 e i decessi/casi totali sia al 31 marzo sia al 30 aprile 2020, in quanto nei grafici 14 e 15 la retta di regressione è caratterizzata rispettivamente da un coefficiente di Pearson maggiore di 0,6 e maggiore di 0,5. Invece non emerge nessuna particolare relazione con i casi totali fino al 28 maggio, poiché in questo caso il coefficiente di Pearson è inferiore a 0,2 (grafico 13).

Detto ciò, di seguito verranno presentati i grafici ottenuti utilizzando il numero di valori di $PM_{10} > 20 \mu g/m^3$.



*Grafico 16: retta di regressione lineare ottenuta per interpolazione del numero di valori di PM_{10} maggiori di $20 \mu g/m^3$ nel periodo 2016-2020 (asse x) e casi totali al 28/05/2020 (asse y), per le 9 province dell'Emilia Romagna. In questa prima analisi si ha un coefficiente $r^2=0,0656$, un **coefficiente di Pearson $r=0,2562$** e $p\text{-value}<0,05$.*

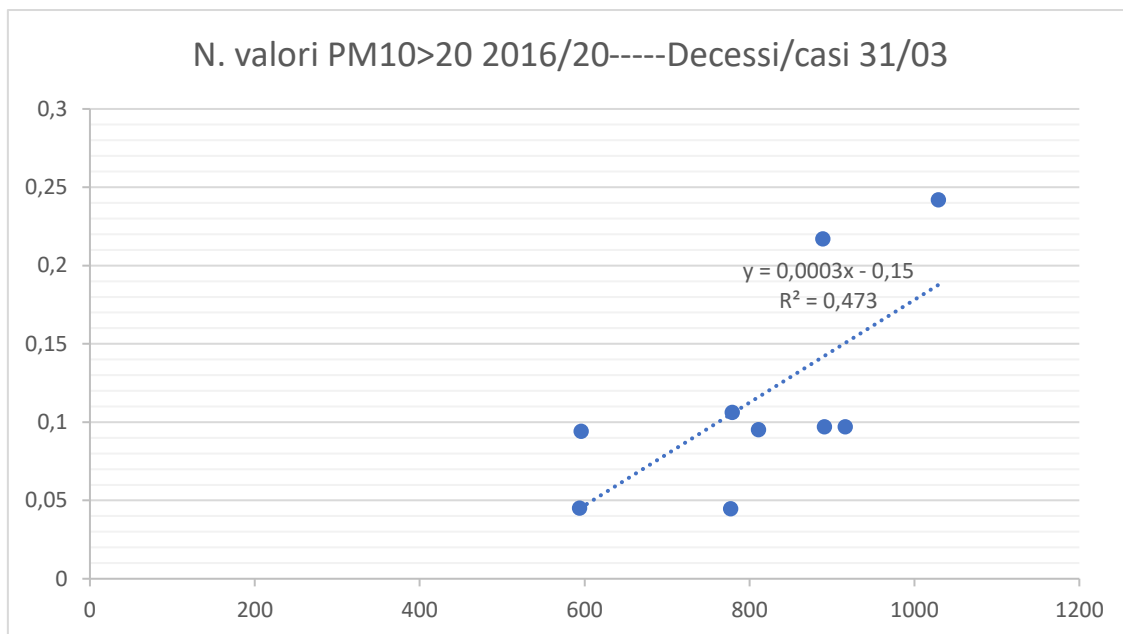


Grafico 17: retta di regressione lineare ottenuta per interpolazione del numero di valori di PM₁₀ maggiori di 20µg/m³ nel periodo 2016-2020 (asse x) e decessi/casi totali al 31/03/2020 (asse y), per le 9 province dell'Emilia Romagna. Il coefficiente r^2 vale 0,4730, il **coefficiente di Pearson r** è 0,6877 e $p\text{-value} < 0,05$.

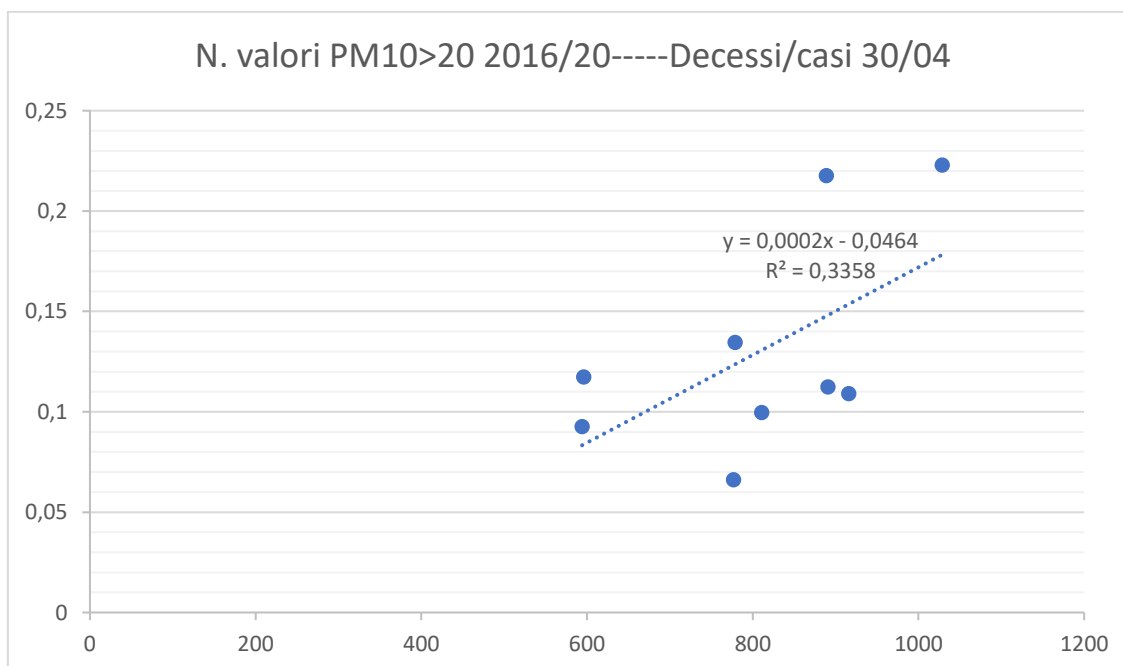


Grafico 18: retta di regressione lineare ottenuta per interpolazione del numero di valori di PM₁₀ maggiori di 20µg/m³ nel periodo 2016-2020 (asse x) e decessi/casi totali al 30/04/2020 (asse y), per le 9 province dell'Emilia Romagna. Il coefficiente r^2 vale 0,3358, il **coefficiente di Pearson r** è 0,5795 e $p\text{-value} < 0,05$.

I risultati appena esposti attraverso i grafici 16,17 e 18, sono analoghi a quelli riportati nei grafici 13,14 e 15, tuttavia la correlazione statistica tra le variabili è ancora più accentuata. Infatti questi ultimi grafici mostrano una buona correlazione tra il numero dei valori di PM_{10} maggiori di $20\mu g/m^3$ e la mortalità del virus (decessi/casi totali), essendo il coefficiente di Pearson dei grafici 17 e 18 risulta più elevato dei corrispettivi grafici 14 e 15.

Di seguito sono riportati gli ultimi tre grafici relativi al PM_{10} che considerano invece il numero di superamenti del valore limite giornaliero di $50\mu g/m^3$, indicato dalla normativa, degli ultimi quattro anni.

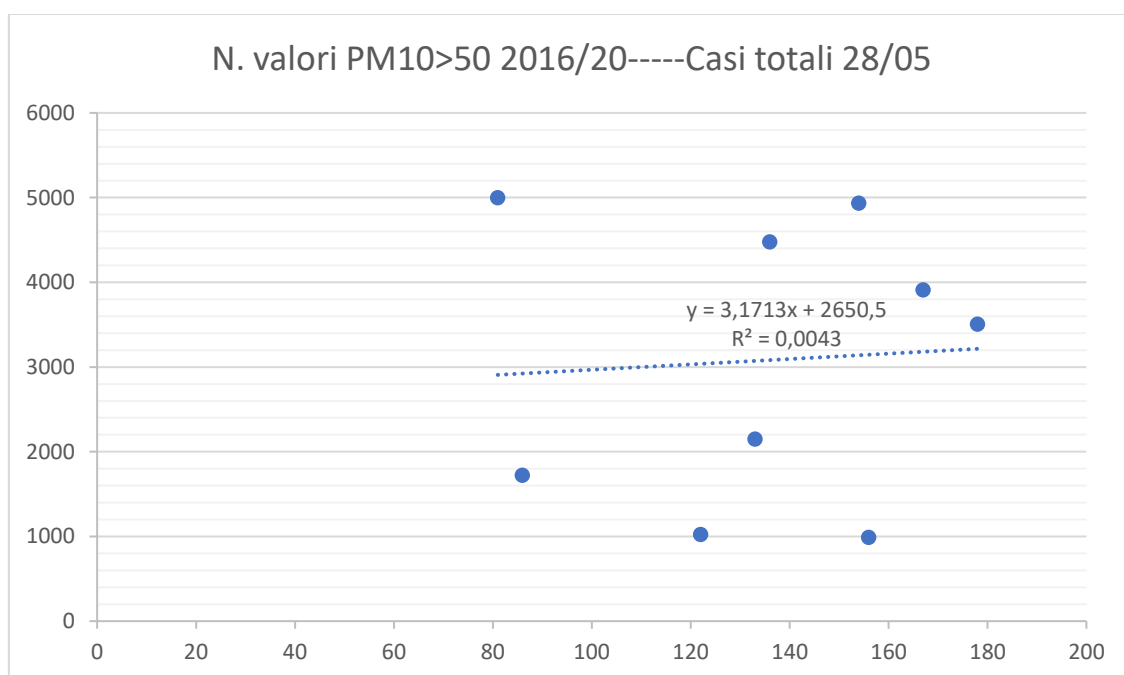


Grafico 19: retta di regressione lineare ottenuta per interpolazione del numero di valori di PM_{10} maggiori di $50\mu g/m^3$ nel periodo 2016-2020 (asse x) e casi totali al 28/05/2020 (asse y), per le 9 provincie dell'Emilia Romagna. In questo caso si ha un coefficiente $r^2=0,0043$, un coefficiente di Pearson $r=0,0659$ e $p\text{-value}<0,05$.

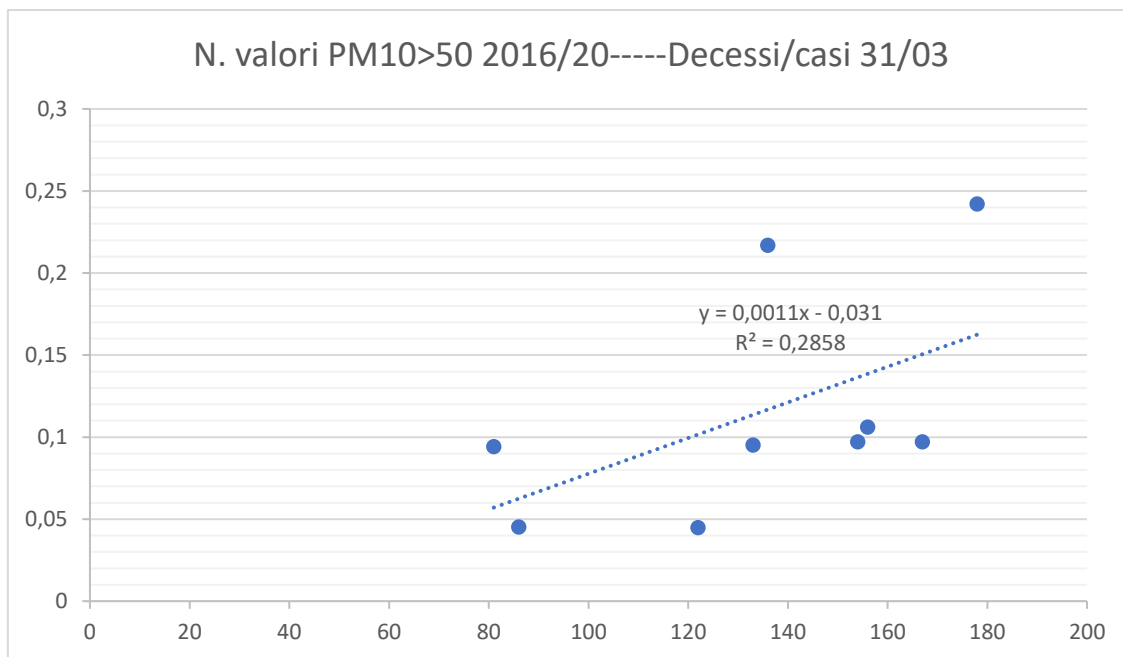


Grafico 20: retta di regressione lineare ottenuta per interpolazione del numero di valori di PM₁₀ maggiori di 50µg/m³ nel periodo 2016-2020 (asse x) e decessi/casi totali al 31/03/2020 (asse y), per le 9 provincie dell'Emilia Romagna. Si ha un coefficiente $r^2=0,2858$, un coefficiente di Pearson $r=0,5346$ e $p\text{-value}<0,05$.

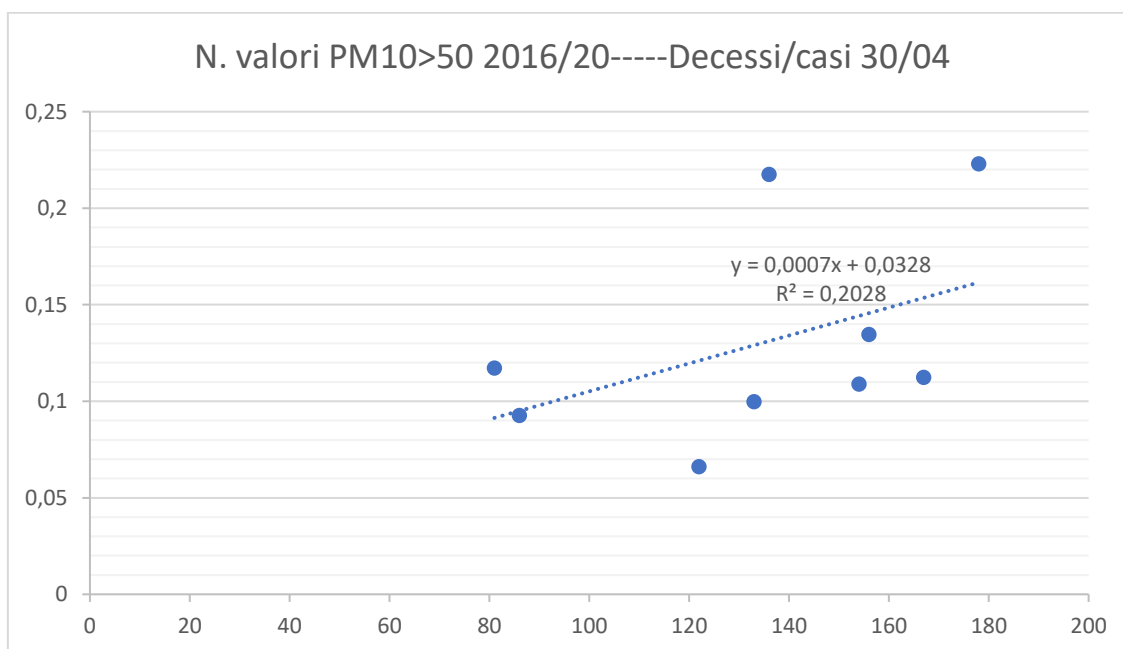
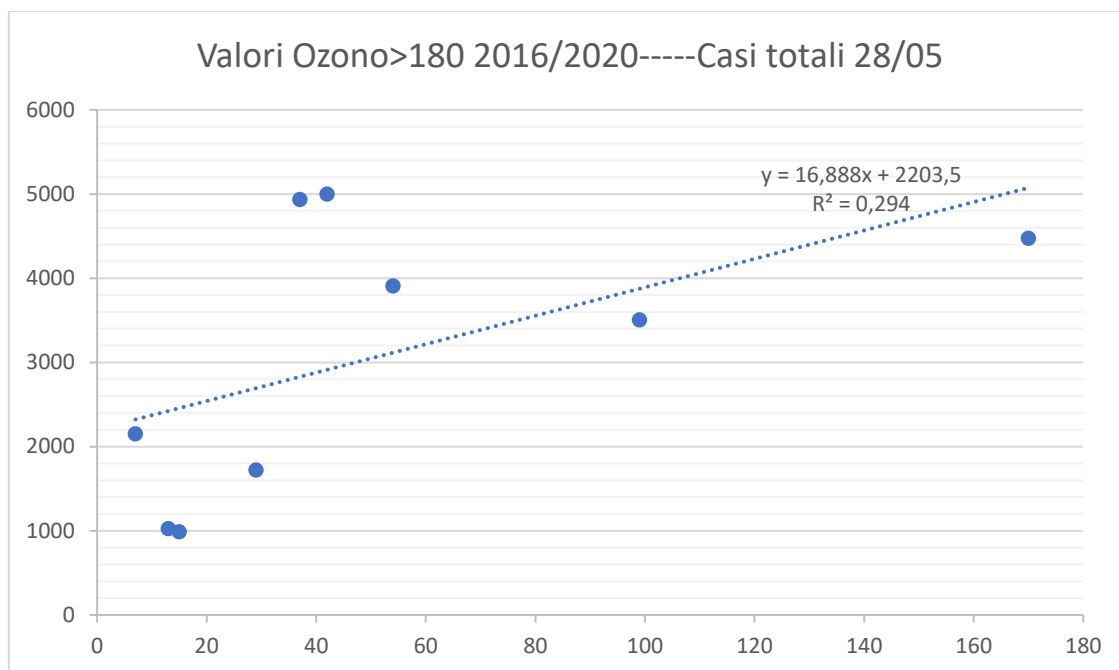


Grafico 21: retta di regressione lineare ottenuta per interpolazione del numero di valori di PM₁₀ maggiori di 50µg/m³ nel periodo 2016-2020 (asse x) e decessi/casi totali al 30/04/2020 (asse y), per le 9 provincie dell'Emilia Romagna. Si ha un coefficiente $r^2=0,2028$, un coefficiente di Pearson $r=0,4503$ e $p\text{-value}<0,05$.

In quest'ultimo caso considerato per il PM₁₀, i risultati trovati sono in linea con i due casi precedenti in cui si è considerata la media e il numero dei valori maggiori di 20µg/m³ del periodo 2016-2020. Infatti i coefficienti di Pearson dei grafici 20 e 21 mostrano una correlazione statistica tra le variabili considerate, anche se più contenuta rispetto ai precedenti grafici 14, 15, 17 e 18. Invece, analogamente ai casi precedenti, dal grafico 19 non emerge nessun risultato interessante per quanto riguarda il numero di superamenti del limite giornaliero e il numero di casi totali di COVID-19.

Infine, verranno di seguito riportati gli ultimi risultati ottenuti per l'Emilia Romagna con l'elaborazione dei dati relativi all'altro inquinante oggetto di studio (O₃).



*Grafico 22: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di O₃>180µg/m³ nel quadriennio 2016-2020 (asse x) e casi totali al 28/05/2020 (asse y) per le 9 province dell'Emilia Romagna. Si ha un coefficiente $r^2=0,2940$, un **coefficiente di Pearson** $r=0,5422$ e $p\text{-value}<0,05$.*

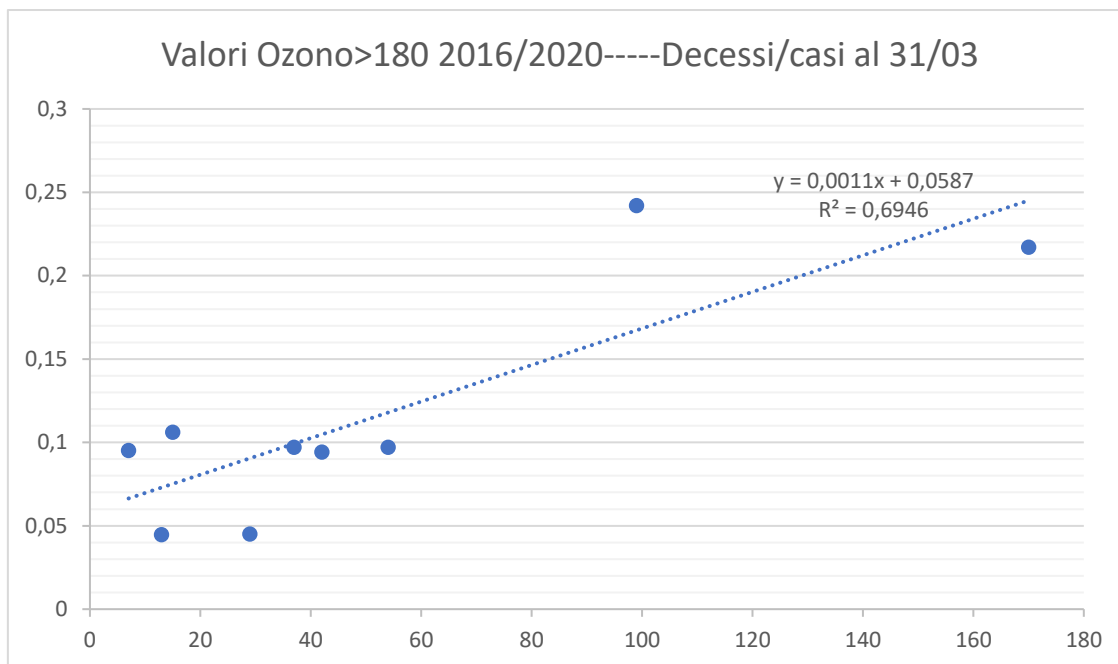


Grafico 23: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $O_3 > 180 \mu\text{g}/\text{m}^3$ nel quadriennio 2016-2020 (asse x) e decessi/casi totali al 31/03/2020 (asse y) per le 9 provincie dell'Emilia Romagna. Si ha un coefficiente $r^2=0,6946$, un coefficiente di Pearson $r=0,8334$ e $p\text{-value}<0,05$.

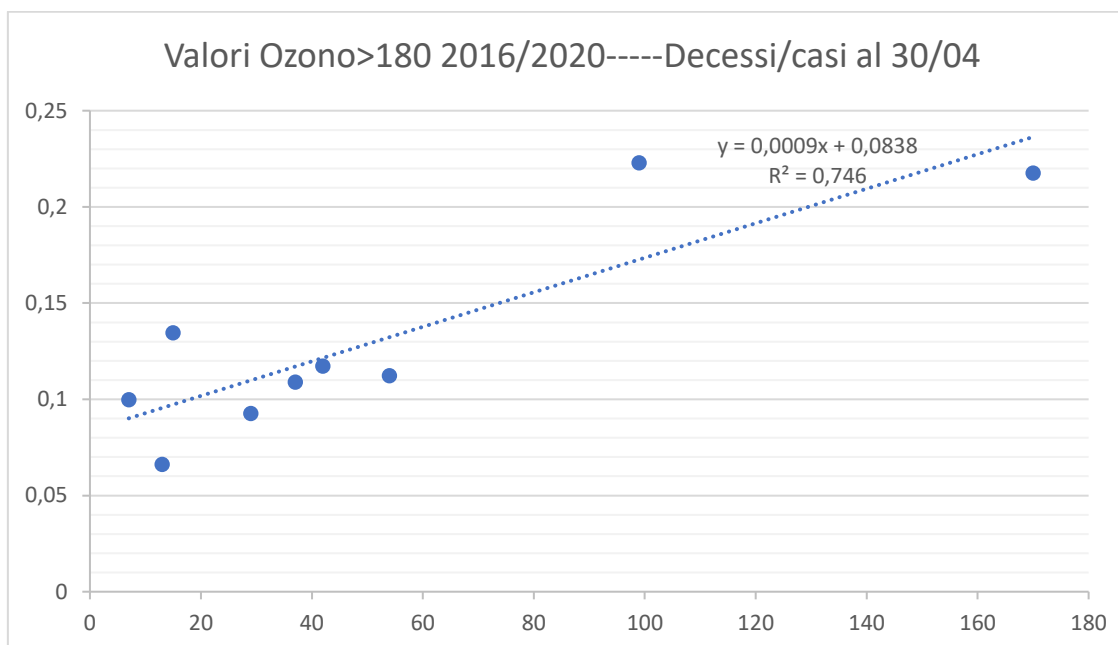


Grafico 24: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $O_3 > 180 \mu\text{g}/\text{m}^3$ nel quadriennio 2016-2020 (asse x) e decessi/casi totali al 30/04/2020 (asse y) per le 9 provincie dell'Emilia Romagna. Si ha un coefficiente $r^2=0,7460$, un coefficiente di Pearson $r=0,8637$ e $p\text{-value}<0,05$.

Dai risultati ottenuti per la regione Emilia Romagna, diversamente dalla regione Lombardia, emerge una discreta proporzionalità tra i valori di PM₁₀ considerati e la mortalità del virus (decessi/casi totali sia al 31/03 che al 30/04), poiché nei grafici 14, 15, 17, 18, 20 e 21 il coefficiente di Pearson risulta sempre almeno maggiore di 0,4 con picchi di quasi 0,7. Invece non emerge nessun risultato degno di nota per quanto riguarda i dati di PM₁₀ presi in considerazione e i casi totali di infezione fino al 28 maggio.

Un'altra importante novità, rispetto alla regione Lombardia, è data dai grafici ottenuti considerando l'O₃: i risultati mostrano una buona correlazione per quanto riguarda il numero di superamenti del valore soglia di 180µg/m³ e il numero di casi totali fino al 28/05 (coefficiente di Pearson maggiore di 0,5); tuttavia i risultati veramente rilevanti sono quelli che si ottengono considerando i decessi/casi sia al 31/03 che al 30/04. In entrambi i grafici 23 e 24, è possibile notare una fortissima correlazione lineare tra le variabili, in quanto il coefficiente di Pearson assume valori maggiori di 0,8 in tutti e due i casi. In conclusione, in base ai risultati ottenuti, sia il PM₁₀ che l'O₃ potrebbero quindi aver avuto un'influenza non trascurabile sulla mortalità del SARS-CoV-2 in Emilia Romagna.

3.3 Riepilogo dei risultati ottenuti

Per una visualizzazione più schematica di tutti i risultati ottenuti dall'elaborazione delle serie di dati di entrambe le regioni, di seguito sono riportate due tabelle che riassumono tutti i risultati precedentemente visualizzati nei grafici.

Tabella 2: riepilogo dei risultati per la regione Lombardia.

Variabili considerate		Coeff. r ²	Coeff. di Pearson (r)	p-value
Media PM ₁₀ 2016-2020	Casi totali al 28/05/20	0,2685	0,5182	< 0,01
Media PM ₁₀ 2016-2020	Decessi/casi al 31/03/20	0,0313	0,1769	< 0,01

Media PM ₁₀ 2016-2020	Decessi/ casi al 30/04/20	0,0773	0,2780	< 0,01
PM ₁₀ > 20µg/m ³	Casi totali al 28/05/20	0,1558	0,3948	< 0,01
PM ₁₀ > 20µg/m ³	Decessi/ casi al 31/03/20	0,0762	0,2760	< 0,01
PM ₁₀ > 20µg/m ³	Decessi/ casi al 30/04/20	0,1218	0,3490	< 0,01
PM ₁₀ > 50µg/m ³	Casi totali al 28/05/20	0,3186	0,5645	< 0,01
PM ₁₀ > 50µg/m ³	Decessi/ casi al 31/03/20	0,0268	0,1638	< 0,01
PM ₁₀ > 50µg/m ³	Decessi/ casi al 30/04/20	0,0874	0,2957	< 0,01
O ₃ > 180µg/m ³	Casi totali al 28/05/20	0,0044	-0.0667	< 0,01
O ₃ > 180µg/m ³	Decessi/ casi al 31/03/20	0,0355	-0.1885	< 0,01

$O_3 > 180\mu g/m^3$	Decessi/ casi al 30/04/20	0,0011	-0.0336	< 0,01
----------------------	---------------------------------	--------	---------	--------

Tabella 3: riepilogo dei risultati per la regione Emilia Romagna

Variabili considerate		Coeff. r^2	Coeff. di Pearson (r)	p-value
Media PM_{10} 2016-2020	Casi totali al 28/05/20	0,0334	0.1829	< 0,01
Media PM_{10} 2016-2020	Decessi/ casi al 31/03/20	0,4029	0,6347	< 0,01
Media PM_{10} 2016-2020	Decessi/ casi al 30/04/20	0,2845	0,5334	< 0,01
$PM_{10} > 20\mu g/m^3$	Casi totali al 28/05/20	0,0656	0,2562	< 0,01
$PM_{10} > 20\mu g/m^3$	Decessi/ casi al 31/03/20	0,4730	0,6877	< 0,01
$PM_{10} > 20\mu g/m^3$	Decessi/ casi al 30/04/20	0,3358	0,5795	< 0,01
$PM_{10} > 50\mu g/m^3$	Casi totali al 28/05/20	0,0043	0,0659	< 0,01

PM ₁₀ > 50µg/m ³	Decessi/ casi al 31/03/20	0,2858	0,5346	< 0,01
PM ₁₀ > 50µg/m ³	Decessi/ casi al 30/04/20	0,2028	0,4503	< 0,01
O ₃ > 180µg/m ³	Casi totali al 28/05/20	0,2940	0,5422	< 0,01
O ₃ > 180µg/m ³	Decessi/ casi al 31/03/20	0,6946	0,8334	< 0,01
O ₃ > 180µg/m ³	Decessi/ casi al 30/04/20	0,7460	0,8637	< 0,01

4. Confronti e commenti

In questo capitolo verranno discussi i principali risultati presentati nel capitolo precedente, anche attraverso il confronto tra i parametri ottenuti per le due regioni separatamente. Inoltre verranno presentati anche i grafici ottenuti attraverso l'elaborazione dei dati di entrambe le regioni congiuntamente. Infine verrà effettuato un paragone anche tra i risultati precedentemente ottenuti elaborando i valori degli inquinanti e i risultati ottenuti elaborando la densità abitativa delle provincie considerate.

In primo luogo si presentano i risultati ottenuti elaborando insieme tutte le 21 provincie di entrambe le regioni, di seguito riportati in forma grafica nello stesso ordine del precedente capitolo. Questi risultati verranno poi utilizzati come un ulteriore confronto con i risultati ottenuti per Lombardia ed Emilia Romagna separatamente.

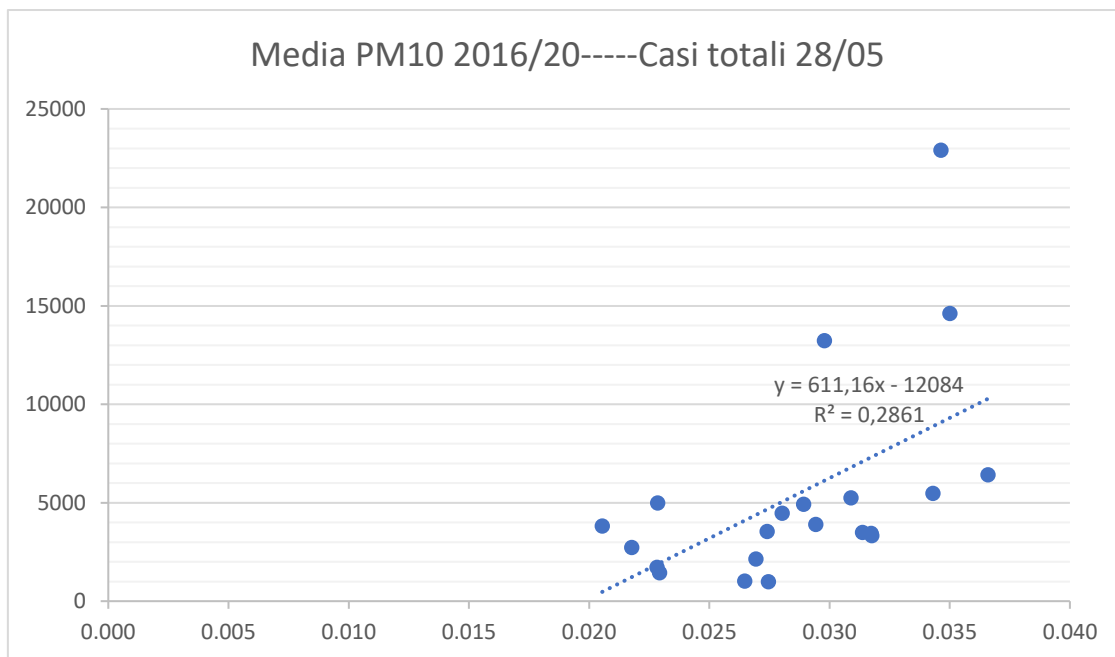


Grafico 25: retta di regressione lineare ottenuta per interpolazione della media di PM_{10} del periodo 2016-2020 (asse x) e i casi totali al 28/05/2020 (asse y), per le 21 provincie di entrambe le regioni. In questo caso si ha un coefficiente $r^2=0,2861$, un **coefficiente di Pearson** $r=0,5349$ e $p\text{-value}<0,05$.

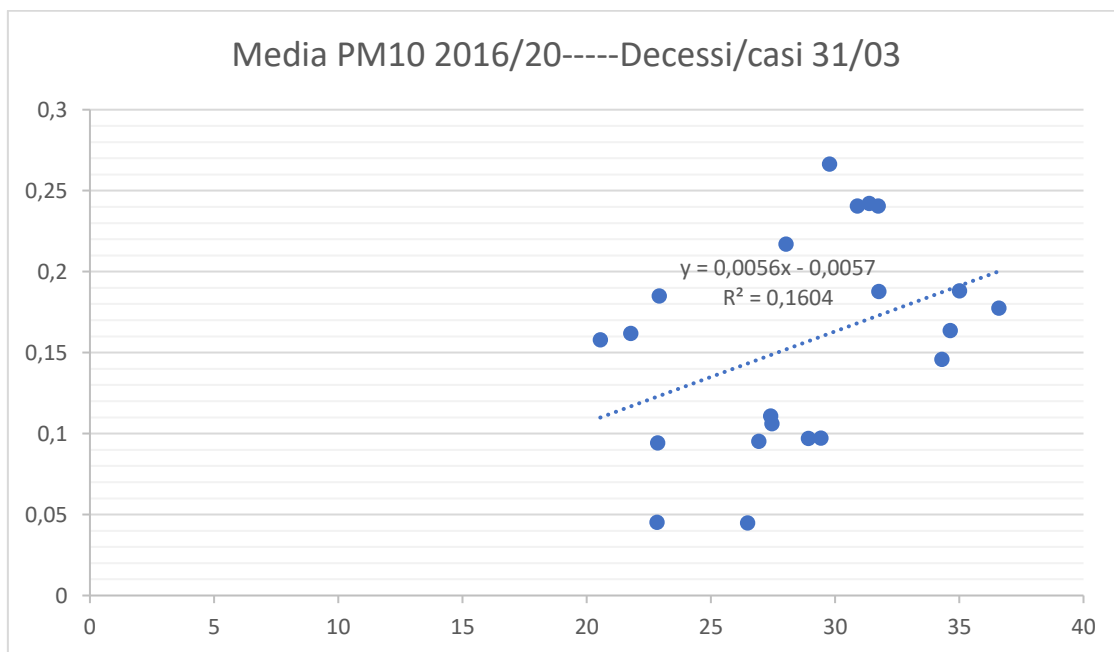


Grafico 26: retta di regressione lineare ottenuta per interpolazione della media di PM_{10} del periodo 2016-2020 (asse x) e i decessi diviso i casi totali al 31 marzo 2020 (asse y), per le 21 provincie di entrambe le regioni. In questo caso si ha un coefficiente $r^2=0,1604$, un **coefficiente di Pearson** $r=0,4004$ e $p\text{-value}<0,05$.

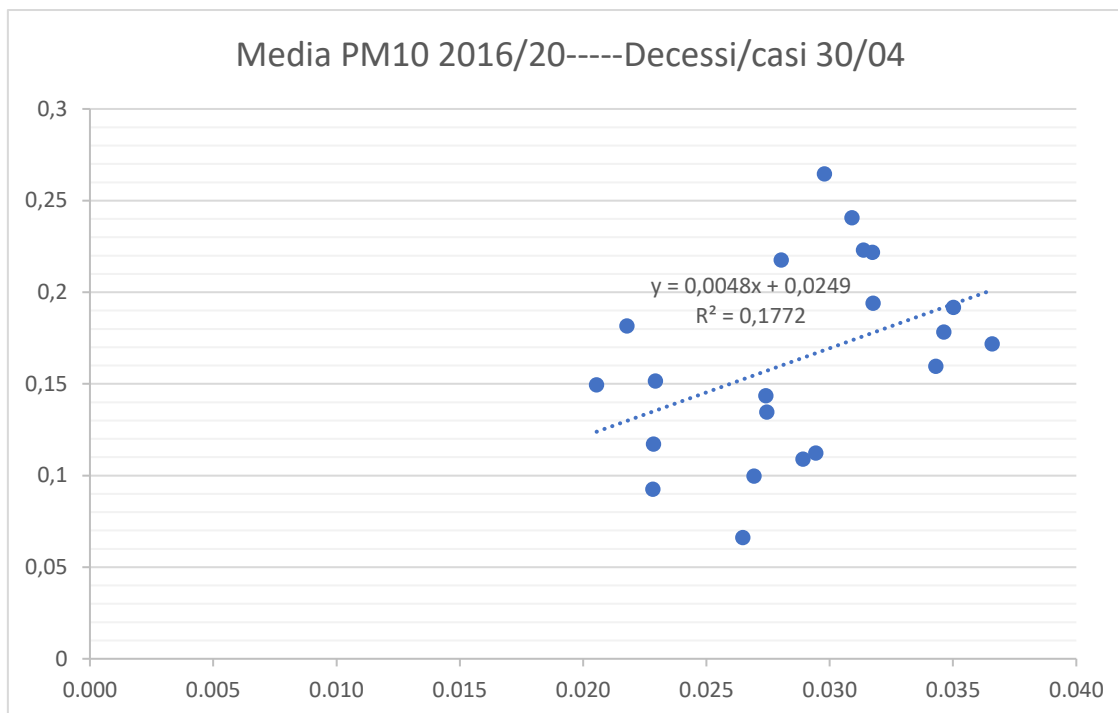


Grafico 27: retta di regressione lineare ottenuta per interpolazione della media di PM_{10} del periodo 2016-2020 (asse x) e i decessi diviso i casi totali al 30 aprile 2020 (asse y), per le 21 provincie di entrambe le regioni. In questo caso si ha un coefficiente $r^2=0,1772$, un coefficiente di Pearson $r=0,4210$ e $p\text{-value}<0,05$.

In tutti e tre i grafici si può notare una relazione tra le variabili degna di nota, poiché il coefficiente di Pearson assume sempre valori maggiori di 0,4. Il valore più elevato del coefficiente si ottiene per il grafico 25 quindi la relazione più forte è con i casi totali, rispetto ai due grafici 26 e 27 che riportano i decessi diviso i casi totali.

Per quanto riguarda il numero di superamenti del valore di $20\mu\text{g}/\text{m}^3$ di PM_{10} nel periodo 2016-2020 e i dati relativi al SARS-CoV-2, i risultati ottenuti sono i seguenti.

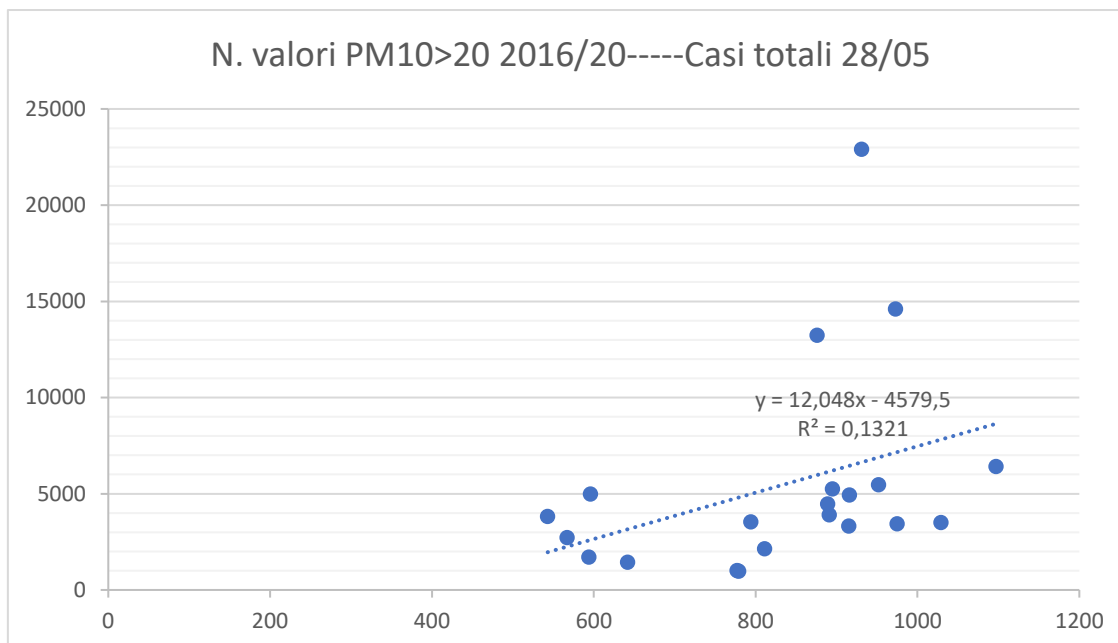


Grafico 28: retta di regressione lineare ottenuta per interpolazione del numero di superamenti del valore di $20\mu\text{g}/\text{m}^3$ di PM_{10} nel periodo 2016-2020 (asse x) e i casi totali al 28/05/2020 (asse y), per le 21 provincie di entrambe le regioni. In questo caso si ha un coefficiente $r^2=0,1321$, un **coefficiente di Pearson $r=0,3634$** e $p\text{-value}<0,05$.

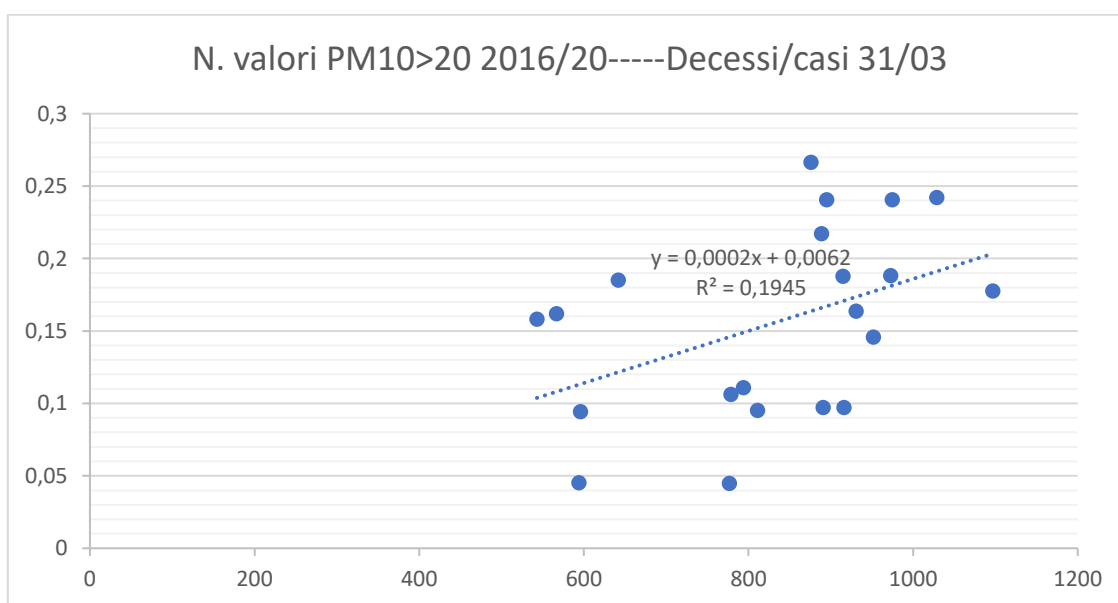
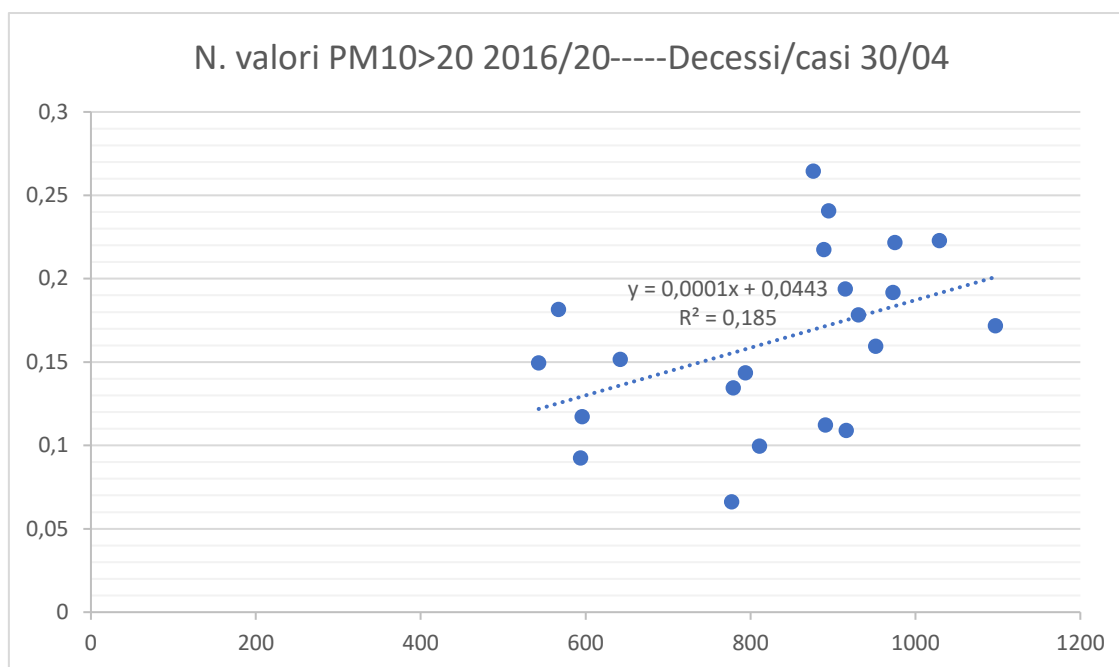


Grafico 29: retta di regressione lineare ottenuta per interpolazione del numero di superamenti del valore di $20\mu\text{g}/\text{m}^3$ di PM_{10} nel periodo 2016-2020 (asse x) e il numero di decessi diviso i casi totali al 31/03/2020 (asse y), per le 21 provincie di entrambe le regioni. In questo caso si ha un coefficiente $r^2=0,1945$, un **coefficiente di Pearson $r=0,4410$** e $p\text{-value}<0,05$.



*Grafico 30: retta di regressione lineare ottenuta per interpolazione del numero di superamenti del valore di $20\mu\text{g}/\text{m}^3$ di PM_{10} nel periodo 2016-2020 (asse x) e il numero di decessi diviso i casi totali al 30/04/2020 (asse y), per le 21 provincie di entrambe le regioni. In questo caso si ha un coefficiente $r^2=0,1850$, un **coefficiente di Pearson $r=0,4301$** e $p\text{-value}<0,05$.*

Questi ultimi tre grafici evidenziano anche in questo caso una correlazione tra le variabili non trascurabile, tuttavia in questo caso i valori del coefficiente di Pearson più elevati si ottengono dalle relazioni con il numero di decessi diviso i casi totali (grafici 29 e 30) rispetto a quella con i casi totali (grafico 28).

Di seguito si riportano le soluzioni ottenute dall'analisi statistica considerando invece il numero di superamenti del valore limite giornaliero di PM_{10} ($50\mu\text{g}/\text{m}^3$), che si sono verificati tra il 2016 e il 2020.

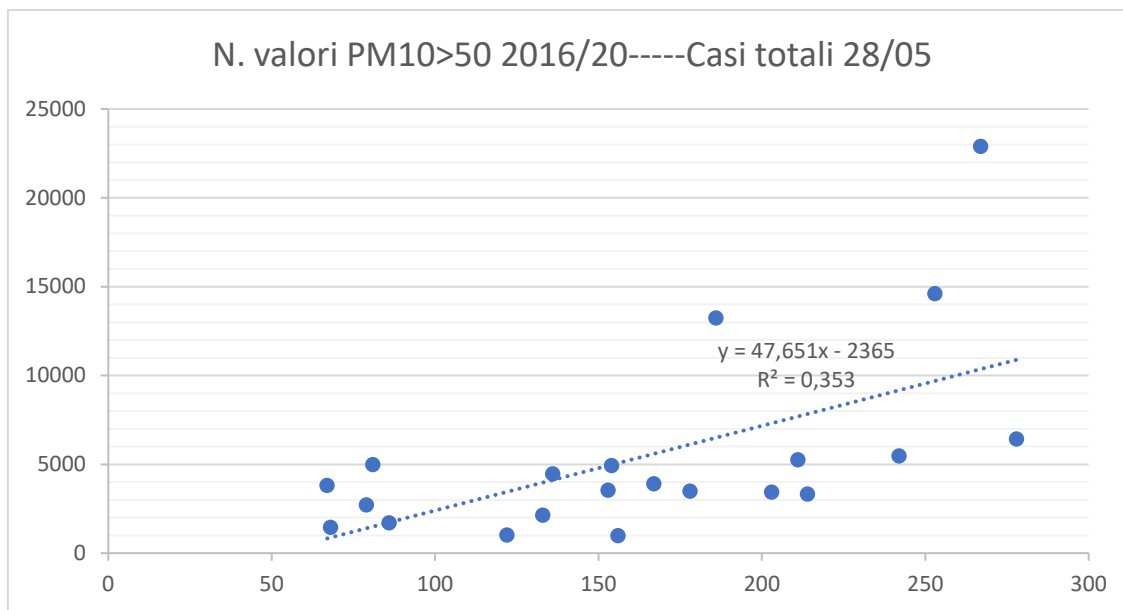


Grafico 31: retta di regressione lineare ottenuta per interpolazione del numero di superamenti del valore limite giornaliero di $50\mu\text{g}/\text{m}^3$ di PM_{10} nel periodo 2016-2020 (asse x) e i casi totali al 28/05/2020 (asse y), per le 21 provincie di entrambe le regioni. In questo caso si ha un coefficiente $r^2=0,3530$, un coefficiente di Pearson $r=0,5942$ e $p\text{-value}<0,05$.

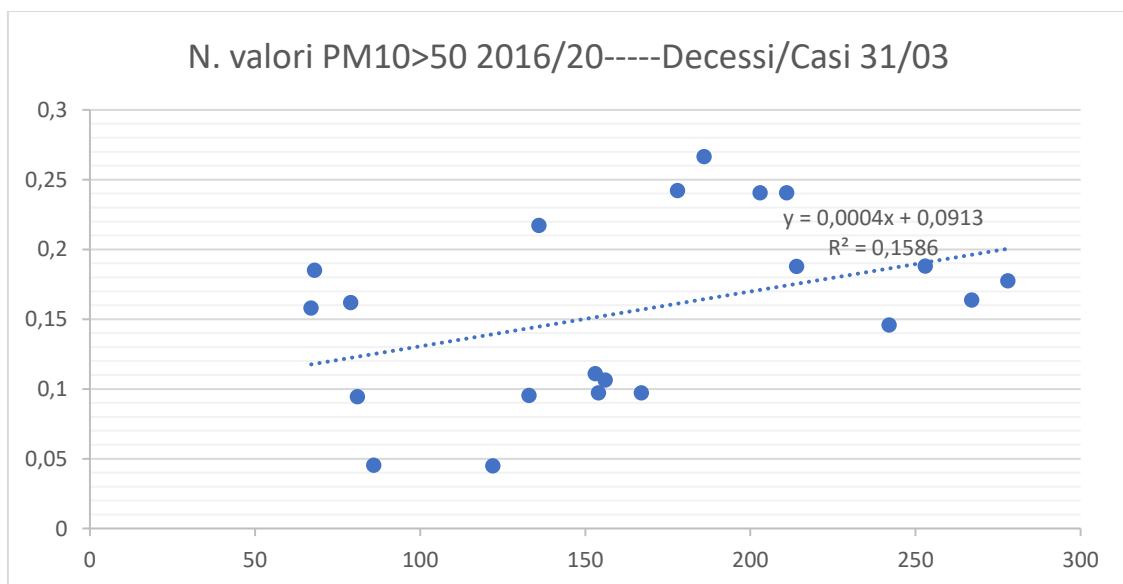


Grafico 32: retta di regressione lineare ottenuta per interpolazione del numero di superamenti del valore limite giornaliero di $50\mu\text{g}/\text{m}^3$ di PM_{10} nel periodo 2016-2020 (asse x) e il numero di decessi diviso i casi totali al 31/03/2020 (asse y), per le 21 provincie di entrambe le regioni. In questo caso si ha un coefficiente $r^2=0,1586$, un coefficiente di Pearson $r=0,3982$ e $p\text{-value}<0,05$.

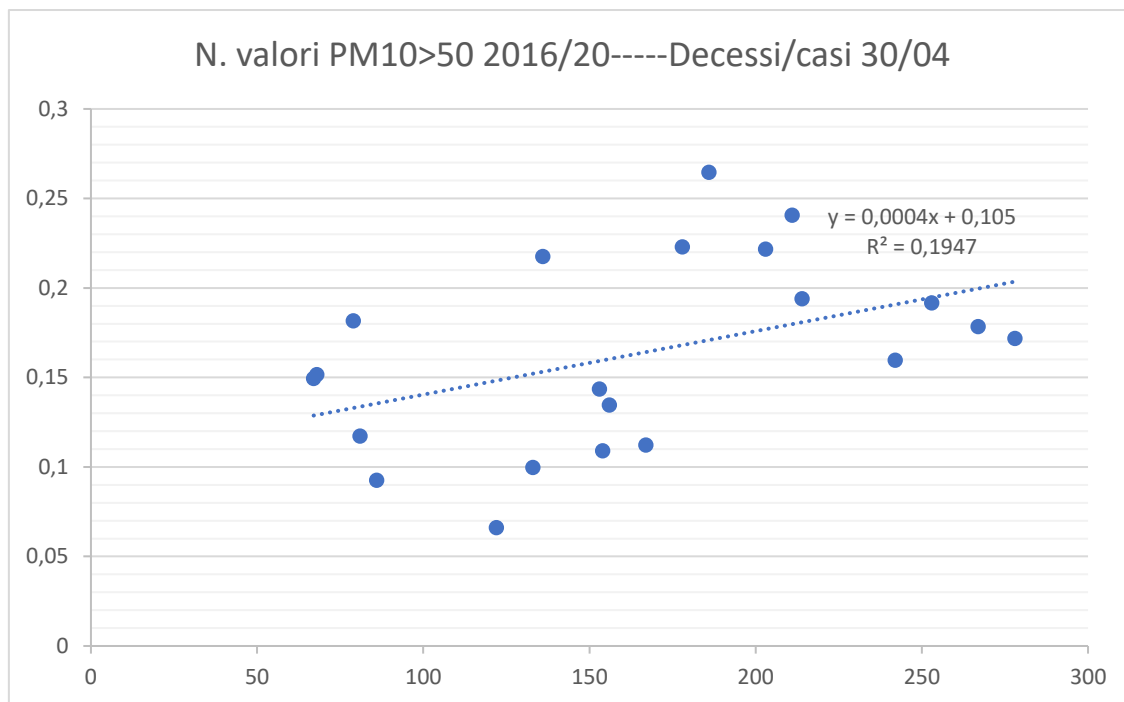


Grafico 33: retta di regressione lineare ottenuta per interpolazione del numero di superamenti del valore limite giornaliero di $50\mu\text{g}/\text{m}^3$ di PM_{10} nel periodo 2016-2020 (asse x) e il numero di decessi diviso i casi totali al 30/04/2020 (asse y), per le 21 provincie di entrambe le regioni. In questo caso si ha un coefficiente $r^2=0,1947$, un coefficiente di Pearson $r=0,4413$ e $p\text{-value}<0,05$.

Si nota anche in questi grafici una buona relazione tra le variabili, in particolare nel grafico 31 che riporta i casi totali al 28 maggio, caratterizzato da un coefficiente di Pearson di quasi 0,6.

Gli ultimi risultati di questa analisi riguardano invece i superamenti del valore di $180\mu\text{g}/\text{m}^3$ dell'ozono troposferico che si sono verificati in tutte le provincie di Lombardia ed Emilia Romagna nel periodo considerato.

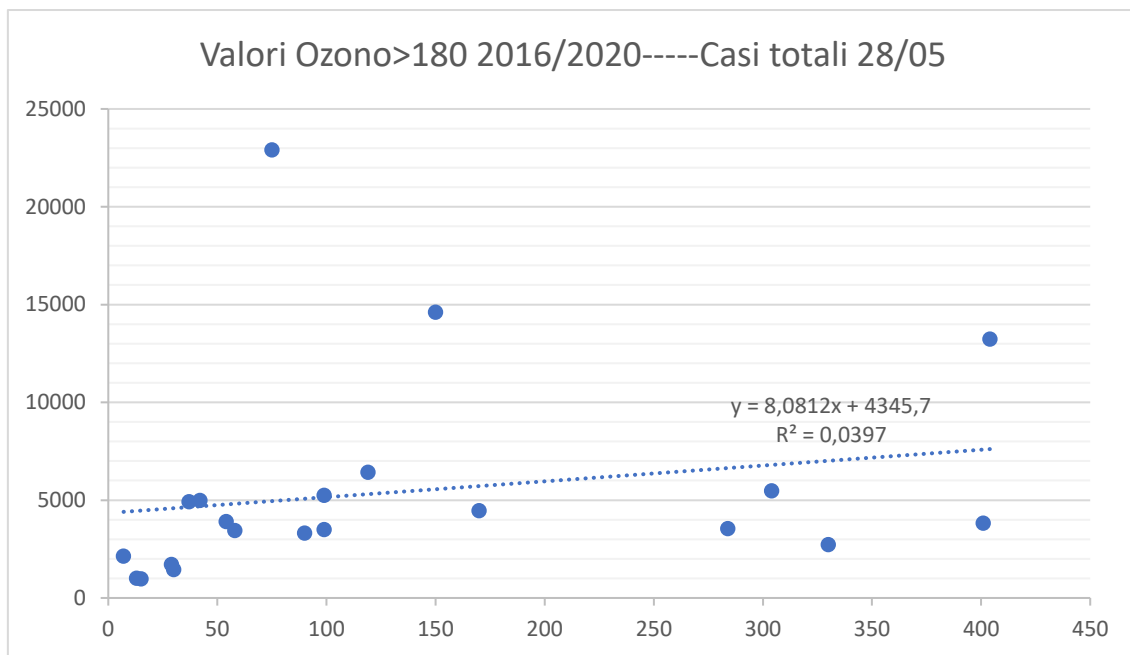


Grafico 34: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $O_3 > 180 \mu\text{g}/\text{m}^3$ nel quadriennio 2016-2020 (asse x) e casi totali al 28/05/2020 (asse y) per le 21 provincie delle due regioni. Si ha un coefficiente $r^2=0,0397$, un **coefficiente di Pearson** $r=0,1992$ e $p\text{-value}<0,05$.

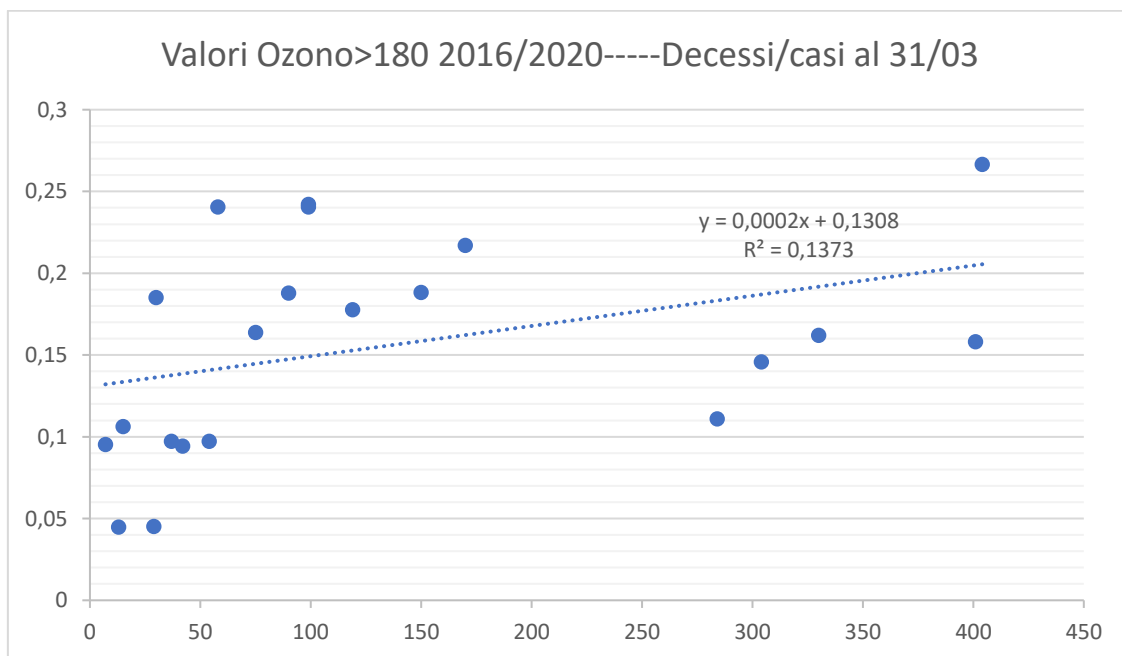
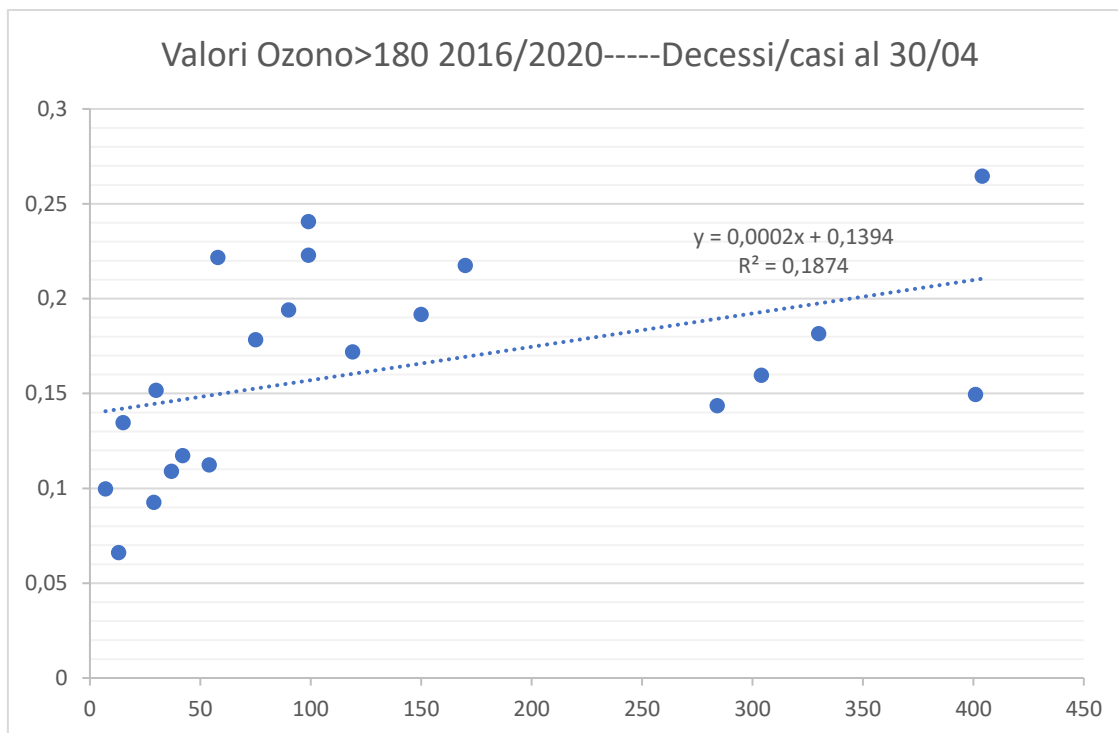


Grafico 35: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $O_3 > 180 \mu\text{g}/\text{m}^3$ nel quadriennio 2016-2020 (asse x) e il numero di decessi diviso i casi totali al 31/03/2020 (asse y) per le 21 provincie delle due regioni. Si ha un coefficiente $r^2=0,1373$, un **coefficiente di Pearson** $r=0,3706$ e $p\text{-value}<0,05$.



*Grafico 36: retta di regressione lineare ottenuta per interpolazione dei punti n. valori di $O_3 > 180 \mu g/m^3$ nel quadriennio 2016-2020 (asse x) e il numero di decessi diviso i casi totali al 30/04/2020 (asse y) per le 21 province delle due regioni. Si ha un coefficiente $r^2 = 0,1874$, un **coefficiente di Pearson $r = 0,4329$** e $p\text{-value} < 0,05$.*

La retta di regressione lineare del grafico 34 è caratterizzata da un coefficiente di Pearson maggiore di zero ma poco elevato (inferiore a 0,2); quindi non emerge nessuna particolare correlazione tra i valori di ozono considerati e i casi totali al 28 maggio. Invece i grafici 35 e 36 riportano rette di regressione caratterizzate entrambe un coefficiente di Pearson prossimo a 0,4 e quindi le variabili prese in considerazione nei due grafici sono legate da una relazione statistica non trascurabile.

4.1 I coefficienti di Pearson a confronto

Di seguito verrà riportato il coefficiente di Pearson di tutte le rette di regressione lineari ottenute per Lombardia, Emilia Romagna e Lombardia + Emilia Romagna congiuntamente. Non verranno invece riportati il coefficiente r^2 (poiché già indicato in tutti i grafici presentati) e il p-value (anch'esso già indicato in tutti i grafici e ampiamente inferiore al 5%, ed in particolare sempre inferiore a 0,01 in tutti i casi).

Tabella 4: riepilogo e confronto tra tutti i coefficienti di Pearson ottenuti.

Variabili considerate		Lombardia	Emilia Romagna	Lombardia + Emilia Romagna
Media PM ₁₀ 2016-2020	Casi totali al 28/05/20	r = 0,5182	r = 0,1829	r = 0,5349
Media PM ₁₀ 2016-2020	Decessi/casi al 31/03/20	r = 0,1769	r = 0,6347	r = 0,4004
Media PM ₁₀ 2016-2020	Decessi/casi al 30/04/20	r = 0,2780	r = 0,5334	r = 0,4210
PM ₁₀ > 20µg/m ³	Casi totali al 28/05/20	r = 0,3948	r = 0,2562	r = 0,3634
PM ₁₀ > 20µg/m ³	Decessi/casi al 31/03/20	r = 0,2760	r = 0,6877	r = 0,4410
PM ₁₀ > 20µg/m ³	Decessi/casi al 30/04/20	r = 0,3490	r = 0,5795	r = 0,4301
PM ₁₀ > 50µg/m ³	Casi totali al 28/05/20	r = 0,5645	r = 0,0659	r = 0,5942

PM ₁₀ > 50µg/m ³	Decessi/ casi al 31/03/20	r = 0,1638	r = 0,5346	r = 0,3982
PM ₁₀ > 50µg/m ³	Decessi/ casi al 30/04/20	r = 0,2957	r = 0,4503	r = 0,4413
O ₃ > 180µg/m ³	Casi totali al 28/05/20	r = -0.0667	r = 0,5422	r = 0,1992
O ₃ > 180µg/m ³	Decessi/ casi al 31/03/20	r = -0.1885	r = 0,8334	r = 0,3706
O ₃ > 180µg/m ³	Decessi/ casi al 30/04/20	r = -0.0336	r = 0,8637	r = 0,4329

Andiamo per ordine: considerando la media del PM₁₀ negli anni 2016-2020 si ha una buona correlazione con i casi totali al 28 maggio per la Lombardia e per Lombardia + Emilia Romagna, mentre risulta debole considerando solo l'Emilia Romagna. Se invece si considerano i decessi/casi totali (sia al 31 marzo che al 30 aprile 2020) si nota una buona correlazione con la media del PM₁₀ per la regione Emilia Romagna e per Lombardia + Emilia Romagna, mentre non si ottiene nessuna relazione per la sola Lombardia.

Prendendo in esame il numero di superamenti di PM₁₀ del valore di 20µg/m³ nel periodo considerato, non si ottiene nessuna particolare relazione con i casi totali al 28/05/20 (il coefficiente di Pearson risulta inferiore a 0,4 in tutti e tre i casi considerati); invece con i decessi/casi totali (sia al 31/03 che al 30/04) si ottiene una debole correlazione per la regione Lombardia, ma questa risulta decisamente più elevata e degna di attenzione per Emilia Romagna e per Lombardia + Emilia Romagna.

Se invece si considerano i superamenti del valore limite di legge di $50\mu\text{g}/\text{m}^3$, si ottiene una buona relazione di proporzionalità con i casi totali al 28 maggio sia per Lombardia che per Lombardia + Emilia Romagna, mentre non si ottiene nessun risultato rilevante per l'Emilia Romagna; con i decessi/casi totali (al 31/03/20 e al 30/04/20) si può osservare una modesta relazione statistica sia per Emilia Romagna che per Lombardia + Emilia Romagna, mentre per la Lombardia la relazione risulta debole.

Infine considerando i valori orari di ozono troposferico maggiori di $180\mu\text{g}/\text{m}^3$ nel periodo 2016-2020, come già introdotto nel capitolo 3, non si ottiene nessun risultato di interesse per la regione Lombardia in nessuno dei tre casi considerati in quanto si hanno coefficienti di Pearson debolmente negativi, in controtendenza con tutti gli altri risultati. Invece per la regione Emilia Romagna si ottiene una buona correlazione con i casi totali al 28 maggio (coefficiente di Pearson maggiore di 0,5) ed una correlazione molto pronunciata con i decessi/casi totali (sia al 31 marzo che al 30 aprile 2020) poiché risulta un coefficiente di Pearson maggiore di 0,8 in entrambi questi ultimi due casi. I risultati ottenuti considerando Lombardia ed Emilia Romagna insieme mostrano una debole correlazione per quanto riguarda i casi totali fino al 28 maggio, mentre si ottiene una modesta relazione di proporzionalità considerando i decessi/casi totali (coefficiente di Pearson circa 0,4 in entrambi i casi).

4.2 Influenza dell'inquinamento atmosferico a confronto con quella della densità abitativa sulla trasmissione del SARS-CoV-2

In aggiunta, sono stati realizzati tre ulteriori grafici ponendo come variabile dipendente il numero dei casi totali al 28 maggio 2020 e come variabile indipendente la densità abitativa (ab/km^2) delle dodici provincie della Lombardia, delle nove provincie dell'Emilia Romagna e delle 21 provincie di entrambe le regioni insieme. Lo scopo di questi ultimi grafici è puramente di confronto, per verificare se l'inquinamento atmosferico possa aver avuto o meno un peso maggiore rispetto alla densità abitativa (e quindi al presunto distanziamento sociale) nella trasmissione del virus, in quanto è ragionevole assumere che nelle provincie con densità abitativa maggiore sia più difficile il rispetto del distanziamento sociale. Di seguito sono riportati i tre grafici.

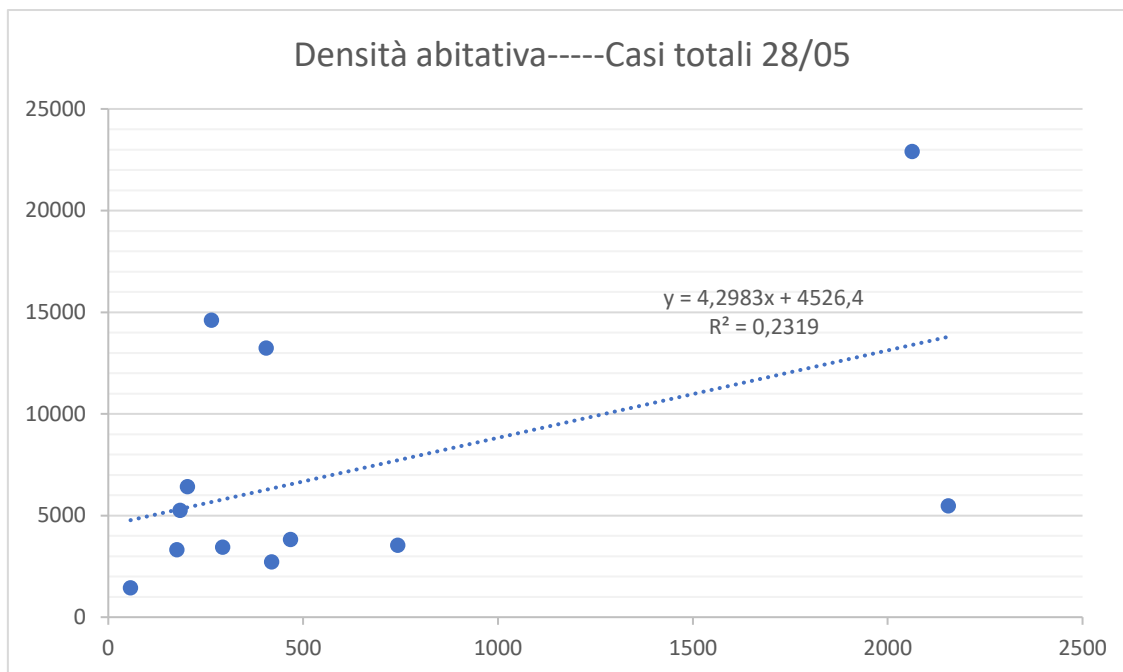


Grafico 37: retta di regressione lineare ottenuta per interpolazione dei punti densità abitativa (ab/km^2) delle 12 provincie lombarde (asse x) e casi totali al 28 maggio 2020 (asse y). Si ha un coefficiente $r^2=0,2319$, un **coefficiente di Pearson** $r=0,4815$ e $p\text{-value}<0,05$.

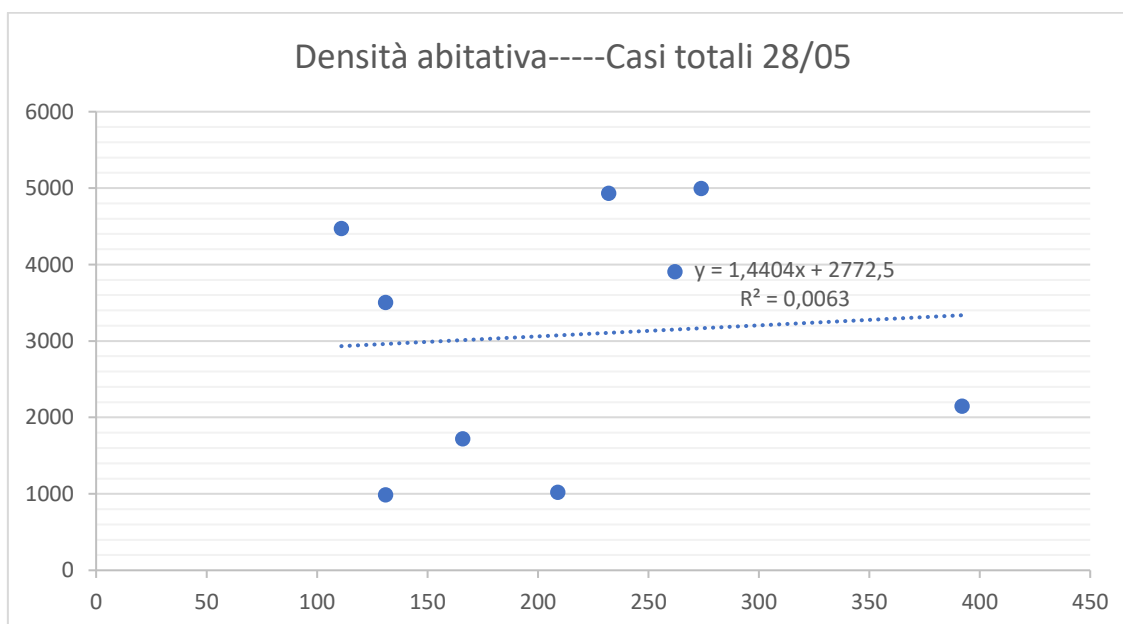
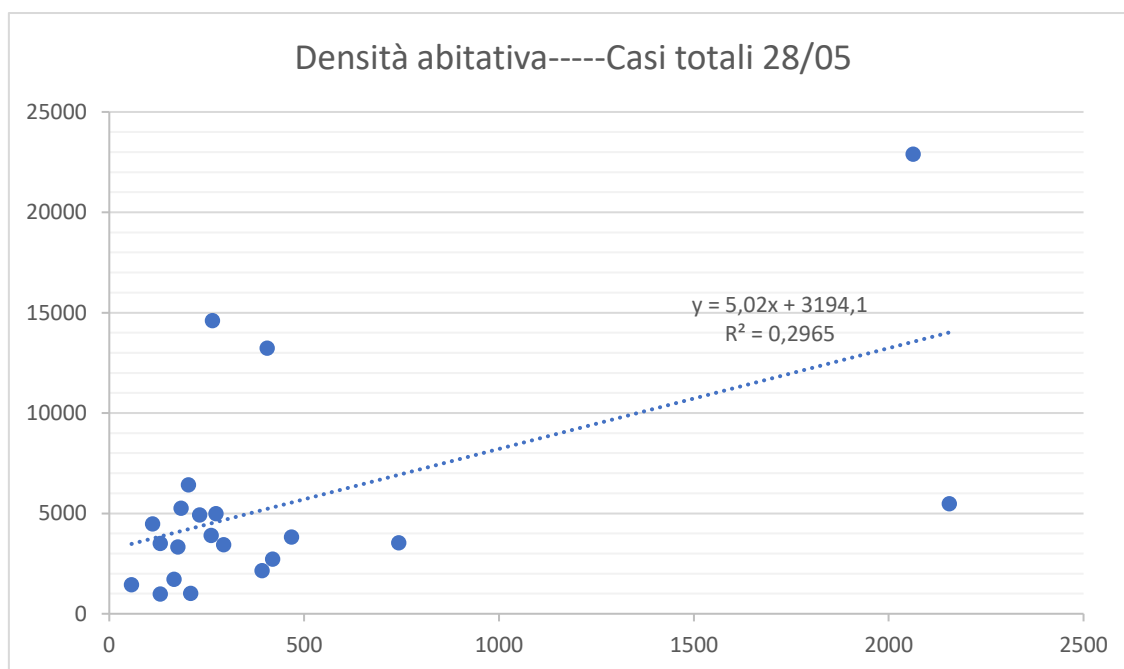


Grafico 38: retta di regressione lineare ottenuta per interpolazione dei punti densità abitativa (ab/km^2) delle 9 provincie dell'Emilia Romagna (asse x) e casi totali al 28 maggio 2020 (asse y). Si ha un coefficiente $r^2=0,0063$, un **coefficiente di Pearson** $r=0,0794$ e $p\text{-value}<0,05$.



*Grafico 39: retta di regressione lineare ottenuta per interpolazione dei punti densità abitativa (ab/km²) delle 21 provincie della Lombardia e dell'Emilia Romagna (asse x) e casi totali al 28 maggio 2020 (asse y). Si ha un coefficiente $r^2=0,2965$, un **coefficiente di Pearson $r=0,5445$** e $p\text{-value}<0,05$.*

Andiamo per ordine: iniziando dalla regione Lombardia risulta che la retta di regressione del grafico 37 è caratterizzata da un coefficiente di Pearson inferiore a quello del grafico 1 e del grafico 7, che riportano entrambi i casi totali al 28 maggio. Per cui il PM₁₀ potrebbe persino aver avuto un'influenza maggiore rispetto alla densità abitativa (e quindi al presunto distanziamento sociale) sulla trasmissione del SARS-CoV-2 nella regione Lombardia.

Considerando invece il grafico 38 della regione Emilia Romagna, risulta che la retta di regressione lineare è caratterizzata da un coefficiente di Pearson minore dei grafici 13 e 16 che riportano la retta di regressione ottenuta considerando i casi totali fino al 28 maggio, e la media del PM₁₀ e il numero di superamenti del valore di 20µg/m³ rispettivamente. Tuttavia in tutti e tre i grafici 13, 16 e 28 e anche nel grafico 19 che riporta i superamenti del valore limite giornaliero per il PM₁₀, il coefficiente di Pearson risulta basso (inferiore a 0,3) quindi si potrebbe escludere a priori un'eventuale relazione statistica tra le variabili considerate in questi grafici. Tuttavia il grafico 22 che riporta la retta di regressione ottenuta considerando l'ozono troposferico e i casi totali fino al 28 maggio 2020 è caratterizzata da un buon coefficiente di Pearson (maggiore di 0,5) e

comunque maggiore di quello del grafico 38, per cui l'O₃ potrebbe aver avuto un impatto maggiore sulla diffusione del COVID-19 nella regione Emilia Romagna.

Infine, il grafico 39 per tutte le provincie di entrambe le regioni è caratterizzato da un coefficiente di Pearson minore del grafico 31 ottenuto considerando i superamenti del valore limite giornaliero per il PM₁₀, ma risulta anche maggiore dei grafici 25 e 28 che riportano rispettivamente la media del PM₁₀ e il numero di superamenti del valore di 20µg/m³ nel periodo 2016-2020. Per cui da questi ultimo confronto non si può affermare se il PM₁₀ abbia avuto o meno un peso maggiore rispetto alla densità abitativa. Inoltre il grafico 34 che riporta il numero di superamenti del valore di 180µg/m³ nel periodo 2016-2020 per l'O₃, risulta caratterizzato da un coefficiente di Pearson inferiore a quello del grafico 39 che riporta la densità abitativa. Per cui anche in questo caso non è possibile affermare se l'O₃ abbia avuto un'influenza maggiore rispetto alla densità abitativa sulla trasmissione del SARS-CoV-2.

5. Conclusione

È già noto da tempo che gli inquinanti contenuti nell'aria danneggiano la salute umana, ed in particolare le vie aeree; quindi i soggetti esposti ad elevati livelli di inquinamento saranno certamente più propensi a sviluppare malattie respiratorie croniche. Inoltre una lunga esposizione a livelli elevati di inquinamento (come accade sia in Lombardia che in Emilia Romagna, dove si verificano frequenti superamenti dei limiti di legge) può incrementare gli effetti negativi sulla salute dovuti a questi inquinanti, in particolare per quanto riguarda il particolato (Tsai, 2019).

Dai risultati ottenuti da questa analisi statistica emerge che sia il PM_{10} che l' O_3 potrebbero aver influenzato la trasmissione e la mortalità del SARS-CoV-2 in Lombardia ed Emilia Romagna. Infatti in alcune regressioni lineari effettuate sono stati ottenuti coefficienti di Pearson elevati, ad indicare una buona correlazione statistica tra le variabili considerate. In particolare, considerando solo i coefficienti di Pearson maggiori di 0,5 che sono stati ottenuti per le due regioni separatamente, si nota che gli inquinanti considerati potrebbero aver avuto una maggiore influenza sulla trasmissione del SARS-CoV-2 in Lombardia mentre potrebbero aver avuto una maggiore influenza sulla mortalità del virus in Emilia Romagna. Inoltre, in entrambe le regioni l'inquinamento atmosferico potrebbe anche aver avuto un'influenza maggiore sulla trasmissione del SARS-CoV-2 rispetto alla densità abitativa.

Per cui, avendo effettuato esclusivamente un'analisi statistica che trascura molte variabili che invece potrebbero aver influenzato i risultati, non è possibile affermare con certezza che l'inquinamento dovuto ad O_3 e PM_{10} abbia certamente influito sulla trasmissione e sulla mortalità del virus. Tuttavia l'evidenza dei risultati di questa tesi suggerisce che alcune delle variabili considerate sono statisticamente correlate tra loro; quindi non si può neanche escludere che l'esposizione cronica ad elevati livelli di inquinamento, negli anni precedenti alla comparsa del SARS-CoV-2, abbiano aggravato gli effetti della pandemia dovuta al COVID-19 in Lombardia ed Emilia Romagna.

6. Bibliografia e sitografia

- World Health Organization: [https://www.who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health)
- Benvenuto, D. G. (2020). The 2019-new coronavirus epidemic: evidence for virus evolution. *Journal of medical virology*, 92(4), 455-459.
- Carugno, M. C. (2017). Temporal trends of PM10 and its impact on mortality in Lombardy, Italy. *Environmental Pollution*, 227, 280-286.
- Chan, J. K. (2020). Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging microbes & infections*, 9(1), 221-236.
- *Coronavirus disease (COVID-19)*. (2020). Tratto da World Health Organization: https://www.who.int/emergencies/diseases/novel-coronavirus-2019?gclid=Cj0KCQjw7ZL6BRCmARIsAH6XFDLqffzc5ljvMJFjjNoKj5UP3xC7SoAL8T-gkkaa94QjT0NfVKGP3UaAIHREALw_wcB
- Cui, Y. Z. (2003). Air pollution and case fatality of SARS in the People's Republic of China: an ecologic study. *Environmental Health*, 2(1), 1-5.
- de Groot, R. J. (2013). Commentary: Middle east respiratory syndrome coronavirus (mers-cov): announcement of the coronavirus study group. *Journal of virology*, 87(14), 7790-7792.
- European Environmental Agency. (2019). *Qualità dell'aria in Europa - Rapporto 2019*.
- Flaxman, S. M.-1. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 10.1038/s41586-020-2405-7.
- Ghebreyesus, T. A. (2020, Marzo 11). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. Tratto da World Health Organization: <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>
- Johnson NPAS, M. J. (2002). Updating the accounts: global mortality of the 1918–1919 “Spanish” influenza pandemic. *Bull Hist Med*, 105-115.
- Kilbourne, E. D. (2006). Influenza pandemics of the 20th century. *Emerging infectious diseases*, 12(1).
- Li, W. S. (2005). Bats are natural reservoirs of SARS-like coronaviruses. *Science*, 310(5748), 676-679.

- Lu, R. Z. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, 395(10224), 565-574.
- Mohd, H. A.-T. (2016). Middle East respiratory syndrome coronavirus (MERS-CoV) origin and animal reservoir. *Virology journal*, 13(1), 1-7.
- Monto, A. &. (1973). Effect of neuraminidase antibody on Hong Kong influenza. *The Lancet*, 301(7804), 623-625.
- Perrino, C. C. (2020). Chemical Composition of PM10 in 16 Urban, Industrial and Background Sites in Italy. *Atmosphere*.
- Ruan, Q. Y. (2020). Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive care medicine*, 46(5), 846-848.
- Tsai, D. R. (2019). Effects of short-and long-term exposures to particulate matter on inflammatory marker levels in the general population. *Environmental Science and Pollution Research*, 26(19), 19697-19704.
- Viboud, C. G. (2005). Multinational impact of the 1968 Hong Kong influenza pandemic: evidence for a smoldering pandemic. *The Journal of infectious diseases*, 192(2), 233-248.
- World Health Organization (2009). Tratto da https://www.who.int/influenza/resources/documents/pandemic_guidance_04_2009/en/
- Wu, A. P. (2020). Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell host & microbe*, 325-328.
- Zhou, P. Y. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *nature*, 579(7798), 270-273.
- Zhou, P. Y. (2020). Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. *BioRxiv*.
- Zhu, N. Z. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*.
- Slides del professor Passerini Giorgio
- Samuel J. Williamson, Fundamentals of Air Pollution, Harlow, Longman Higher Education, 1973

- IARC Monograph on the evaluation of carcinogenic risk to human; volume 109 – Outdoor air pollution, 2015 <http://monographs.iarc.fr/ENG/Monographs/vol109/index.php>
- <https://www.issalute.it/index.php/la-salute-dalla-a-alla-z-menu/p/pm10-particolato-atmosferico-o-polveri-sottili#effetti-sulla-salute>
- <https://www.epicentro.iss.it/ambiente/ozonoUe>
- Decreto Legislativo 13 agosto 2010, n.155: “Attuazione della direttiva 2008/50/CE relativa alla qualità dell'aria ambiente e per un'aria più pulita in Europa”
- <https://www.sciencedirect.com/science/article/abs/pii/S0269749116321315>
- [https://www.who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health)
- Leslie Collier, Albert Balows e Max Sussman, Virology, in Brian H Mahy e Leslie Collier (a cura di), Topley and Wilson's Microbiology and Microbial Infections, vol. 1, 9ª ed., Arnold, 1998
- Nigel J. Dimmock, Andrew J. Easton e Keith N. Leppard, Introduction to Modern Virology, VI, Blackwell Publishing, 2007
- Dorothy H. Crawford, Viruses: A Very Short Introduction, Oxford University Press, 2011
- https://www.who.int/influenza/resources/documents/pandemic_guidance_04_2009/en/
- https://web.archive.org/web/20110513194204/http://www.who.int/csr/disease/influenza/PIP_Guidance09.pdf
- <https://www.who.int/bulletin/volumes/89/7/11-088815/en/>
- <https://web.archive.org/web/20200311214147/https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>
- https://web.archive.org/web/20110513194204/http://www.who.int/csr/disease/influenza/PIP_Guidance09.pdf
- <https://www.epicentro.iss.it/passi/storiePandemia>
- https://www.globalsecurity.org/security/ops/hsc-scen-3_pandemic-1918.htm
- https://www.globalsecurity.org/security/ops/hsc-scen-3_pandemic-1957.htm
- https://www.globalsecurity.org/security/ops/hsc-scen-3_pandemic-1968.htm
- <https://www.cdc.gov/flu/pandemic-resources/1918-pandemic-h1n1.html>
- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3291411/#R16>
- <https://www.cdc.gov/flu/pandemic-resources/2009-h1n1-pandemic.html>
- <https://www.epicentro.iss.it/influenza/PandemiaH1N1>
- <https://www.who.int/emergencies/mers-cov/en/>

- <http://www.salute.gov.it/portale/nuovocoronavirus/dettaglioFaqNuovoCoronavirus.jsp?lingua=italiano&id=228>
- <https://www.epicentro.iss.it/coronavirus/sintomi-diagnosi>
- https://www.arpae.it/dettaglio_generale.asp?id=3882&idlivello=2073
- <https://www.arpalombardia.it/Pages/Aria/Rete-di-rilevamento/Criteri-di-rilevamento.aspx>
- <https://www.arpalombardia.it/Pages/Aria/Rete-di-rilevamento/Zonizzazione.aspx?firstlevel=Rete%20di%20rilevamento>
- https://www.arpae.it/dettaglio_generale.asp?id=2892&idlivello=846
- <http://www.protezionecivile.gov.it/attivita-rischi/rischio-sanitario/emergenze/coronavirus>
- James Stock, Mark Watson, *Introduzione all'econometria*, Milano, Pearson Education, 2005, p. 121
- Draper, N.R. and Smith, H. (1998). *Applied Regression Analysis*. Wiley-Interscience
- Everitt, B.S. (2002). *Cambridge Dictionary of Statistics* (2nd Edition)
- Nagelkerke, Nico J.D. (1992) Maximum Likelihood Estimation of Functional Relationships, Pays-Bas, *Lecture Notes in Statistics*, Volume 69, 110p
- Luigi Fabbri, *Statistica multivariata (analisi esplorativa dei dati)*. 1997, McGrawHill
- Ronald L. Wasserstein e Nicole A. Lazar, The ASA's Statement on p-Values: Context, Process, and Purpose, in *The American Statistician*, vol. 70, n. 2, 7 marzo 2016, pp. 129–133
- Bhaskar Bhattacharya e DeSale Habtzghi, Median of the p value under the alternative hypothesis, in *The American Statistician*, vol. 56, n. 3, 2002, pp. 202–6
- R. Wetzels, D. Matzke, M. D. Lee, J. N. Rouder, G. J. Iverson e E. -J. Wagenmakers, Statistical Evidence in Experimental Psychology: An Empirical Comparison Using 855 t Tests, in *Perspectives on Psychological Science*, vol. 6, n. 3, 2011, pp. 291–298
- E. Babble, *The practice of social research*, 11^a ed., Belmont, Thomson Wadsworth, 2007
- James Stock, Mark Watson, *Introduzione all'econometria*, Milano, Pearson Education, 2005, p. 95
- D.C. Boes, F.A. Graybill, A.M. Mood (1988), *Introduzione alla Statistica*, McGraw-Hill Libri Italia
- N.R. Draper, H. Smith (1998), *Applied Regression Analysis*, Wiley Series in Probability and Statistics
- John R. Taylor, (1999) *Introduzione all'analisi degli errori*, Zanichelli editore

- <https://www.britannica.com/science/planetary-boundary-layer> Retrieved on 2020-06-28
- https://web.archive.org/web/20051215040805/http://www.euro.who.int/document/aicq/7_3particulate-matter.pdf
- <https://www.issalute.it/index.php/saluteaz-saz/p/676-pm10-particolato-atmosferico-opolveri-sottili#link-approfondimento>
- <https://www.arpalombardia.it/Pages/Aria/Inquinanti.aspx>
- <https://www.arpalombardia.it/sites/DocumentCenter/Documents/ARIA-Inquinanti/InformativaOzono.pdf>
- <https://www.regione.lombardia.it/wps/portal/istituzionale/HP/DettaglioRedazionale/servizi-e-informazioni/cittadini/salute-e-prevenzione/Sicurezza-negli-ambienti-di-vita-e-di-lavoro/inquinamento-atmosferico/inquinamento-atmosferico/>
- https://www.isprambiente.gov.it/files2017/pubblicazioni/stato-ambiente/rau-2017/5_Qualita%20dellaria.pdf
- <https://www.arpae.it/motap/ozono/ozono.htm>
- <https://www.eea.europa.eu/it/publications/92-828-3351-8/page005.html>
- <https://acp.copernicus.org/articles/14/4895/2014/acp-14-4895-2014.pdf>
- http://www.esa.int/Applications/Observing_the_Earth/Envisat/Global_air_pollution_map_produced_by_Envisat_s_SCIAMACHY
- https://www.epicentro.iss.it/focus/h1n1/pdf/tabella_Oms_def.pdf
- <https://www.arpalombardia.it/Pages/Aria/Da-sapere.aspx>
- https://www.arpae.it/dettaglio_generale.asp?id=3882&idlivello=2073
- <https://www.arpalombardia.it/Pages/Aria/Rete-di-rilevamento/Zonizzazione.aspx?firstlevel=Rete%20di%20rilevamento>