



## IV GIORNATA DI CARATTERIZZAZIONE CHIMICA DEL PARTICOLATO ATMOSFERICO

Terni, 22-22 Novembre 2022

### Variazioni della composizione del particolato atmosferico durante il 2020: un approccio con una tecnica di machine learning

Autori: Fabiana Scotto<sup>1</sup>, Andrea Algieri<sup>2</sup>, Dimitri Bacco<sup>1</sup>, Arianna Trentini<sup>1</sup>

<sup>1</sup>Agenzia regionale per la prevenzione, l'ambiente e l'energia dell'Emilia-Romagna (Arpae)

<sup>2</sup>Lombardia - Agenzia regionale per la protezione dell'ambiente della Lombardia (ARPAL)

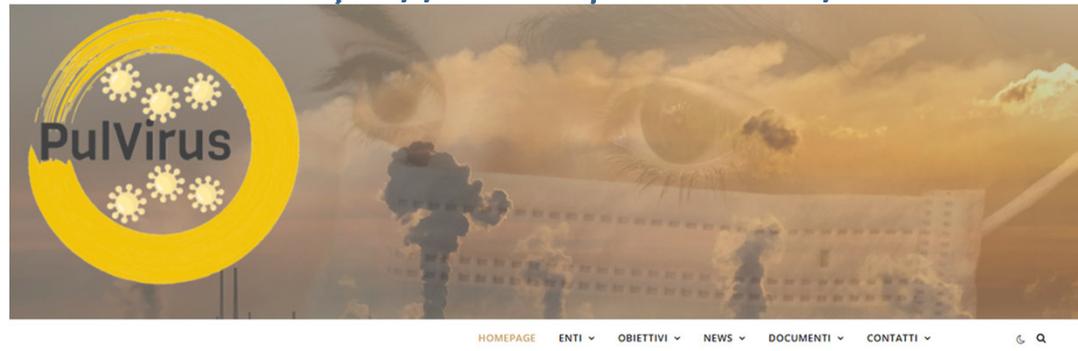


SCUOLA DI ALTA FORMAZIONE AMBIENTALE



## Il progetto Pulvirus

<https://www.pulvirus.it/>



### Progetto Pulvirus

**C**he la "polvere", il particolato atmosferico (PM), fosse un problema serio, soprattutto in Pianura Padana, è noto almeno da quando le reti di misura dell'inquinamento atmosferico hanno cominciato a produrre dati di qualità sistematicamente, ormai son vent'anni, di PM10, di PM2.5 e dei suoi componenti inorganici ed organici. Se il particolato atmosferico, responsabile di morti premature e di problemi respiratori e cardiocircolatori acuti, sia anche il "vettore", il supporto del virus SARS-COV2 responsabile dell'espandersi dell'epidemia nelle Regioni del Nord Italia è una domanda circolata nei giorni neri. Una domanda troppo importante con implicazioni così spaventose non poteva lasciare indifferente la comunità scientifica e non poteva non suggerire la cautela.

**P**ulvirus intende fornire risposte ragionate e scientificamente approfondite. È il prodotto di un'alleanza fra le principali Istituzioni tecnico-scientifiche del Paese che agiscono a livello nazionale e a livello locale e che hanno la responsabilità di informare i decisori politici e l'opinione pubblica. Ogni parola, ogni presa di posizione, deve essere quindi soppesata e condivisa perché può indirizzare un provvedimento dell'autorità o formare l'opinione del pubblico.

**F**NEA, SNPA, ISS metteranno in comune le proprie competenze ed i propri "attrezzi di lavoro" (reti di misura, dati, modelli, strategie di sorveglianza epidemiologica, infrastrutture di calcolo, capacità analitica), valorizzando le collaborazioni già in corso con altri soggetti pubblici e coordinandosi con alcune iniziative, come LIFE PREPAIR e lo sviluppo di Space Economy-Mirror Copernicus, che vedono SNPA protagonista.

## Le domande da cui siamo partiti

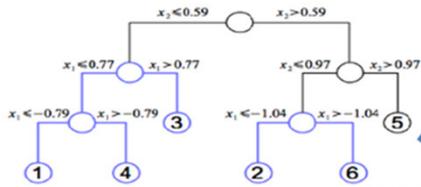
La composizione del particolato è stata influenzata dalle azioni intraprese per limitare il diffondersi della pandemia?

Eventuali variazioni osservate potrebbero essere imputabili alla meteorologia?

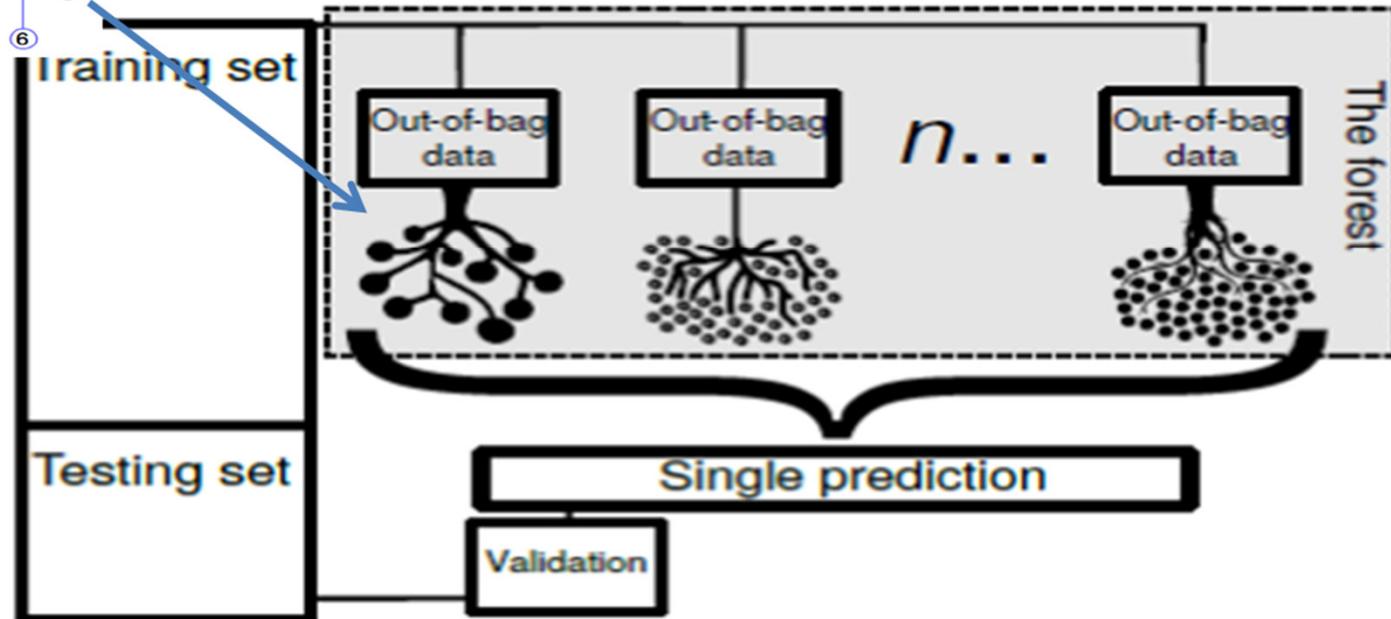
## La tecnica RF

Tramite un modello di machine learning (Random Forest) si sono stimate le concentrazioni di inquinanti attese in base alle condizioni meteorologiche verificatesi nel 2020 (stima “Business As Usual” - BAU) e in assenza delle misure restrittive messe in atto.

La **differenza** tra le concentrazioni previste e quelle osservate si è considerata imputabile ai **provvedimenti** presi per limitare il diffondersi della pandemia.



La tecnica Random Forest (RF), introdotta da Breiman (2001).



## La tecnica RF: vantaggi

- Tecnica non parametrica di machine learning.
- Ha il grande vantaggio di non dover fare assunzioni come la normalità del campione, l'omoschedasticità, l'indipendenza, l'aderenza ad altre rigorose ipotesi parametriche e l'attenta gestione degli effetti di interazione che sono invece richieste per i modelli statistici.

## La tecnica RF: svantaggi

È in grado di stimare una componente di trend sul campione di dati utilizzato per addestrare il modello, ma **non è in grado di estrapolare il trend** per un periodo di dati successivo (o antecedente) il periodo di addestramento.

L'incertezza del modello e delle stime non è espressa secondo i classici "intervalli di confidenza"

## Previsione Business As Usual

Periodo di **addestramento** del modello: 1° gennaio 2013 - 31 dicembre 2019

Periodo di **validazione** del modello: 1° gennaio 2020 - 14 febbraio 2020

**Stime "Business As Usual"**: 15 febbraio 2020 - 31 dicembre 2020

## Previsione Business As Usual

Non tutte le specie analizzate hanno dato risultati sufficientemente buoni: si è deciso di considerare solo le specie con un fitting del modello per cui  $R^2 \simeq 0.5$  e per cui la differenza tra osservati e predetti nel 2020 fosse superiore alle differenze tra osservati predetti ottenute per gli anni della serie storica usata per addestrare il modello

## Siti e specie chimiche analizzate

Sono stati studiati siti per cui fosse disponibile una serie storica di caratterizzazione chimica del particolato sufficientemente lunga:

- Milano - Urban Background
  - Milano - Traffico
  - Bologna - Urban Background
  - Parma - Urban Background
  - Rimini - Urban Background
  - S. Pietro Capofiume - Rural Background
- } PM10: EC, OC, ioni (nitrato, solfato, ammonio, ...), levoglucosano, elementi in tracce (Al, Cu, Fe, ...)
- } PM2.5: EC, OC, ioni (nitrato, solfato, ammonio, ...), levoglucosano, elementi in tracce (Fe, V, Sb, ...)

Sono state studiate inoltre serie storiche di **BC** (**BC<sub>ff</sub>** e **BC<sub>bb</sub>**) e **NH<sub>3</sub>** nei siti in cui erano disponibili.

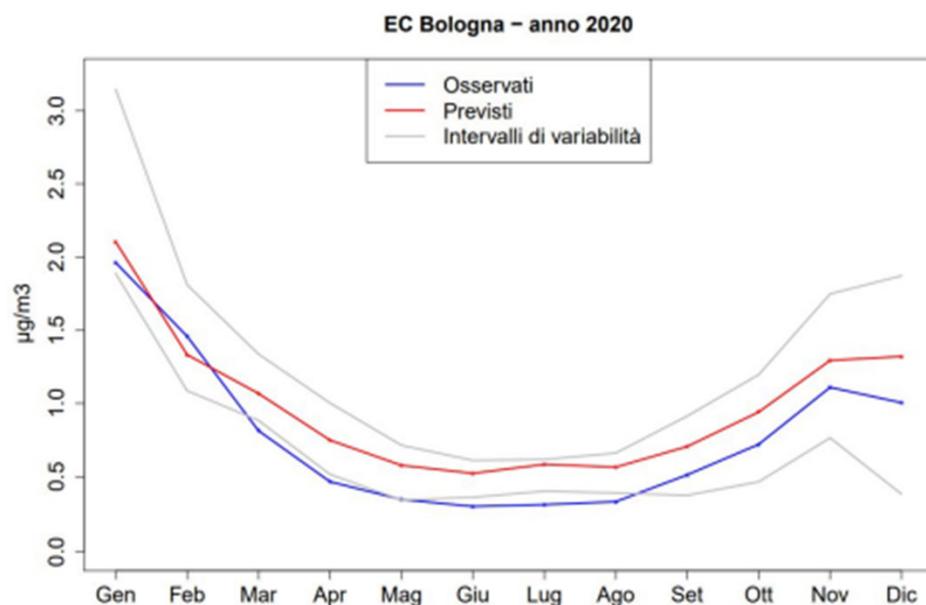
## Le variabili esplicative utilizzate

Le **variabili meteorologiche** selezionate:

- Pressione
- Temperatura media (2m)
- Precipitazione cumulata
- Precipitazione cumulata del giorno precedente
- Umidità relativa
- Irradianza solare netta
- Altezza dello strato limite planetario (minimo e massimo giornaliero)
- Direzione e velocità del vento

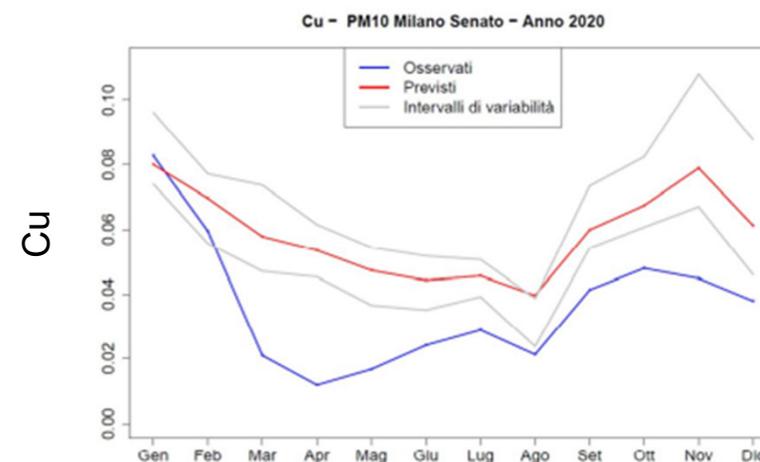
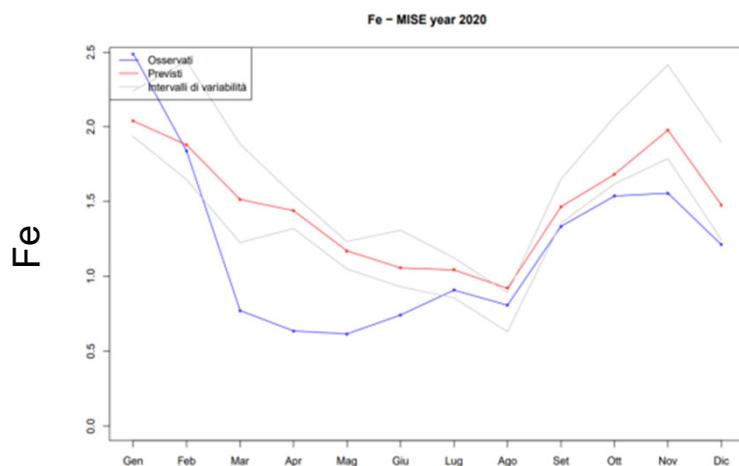
Inoltre sono stati utilizzati un **termine di tendenza** sotto forma di data Unix (numero di secondi dal 1° gennaio 1970), e **due termini stagionali** sotto forma di giorno giuliano (giorno dell'anno) e giorno della settimana.

## Risultati: Carbonio Elementare (EC)



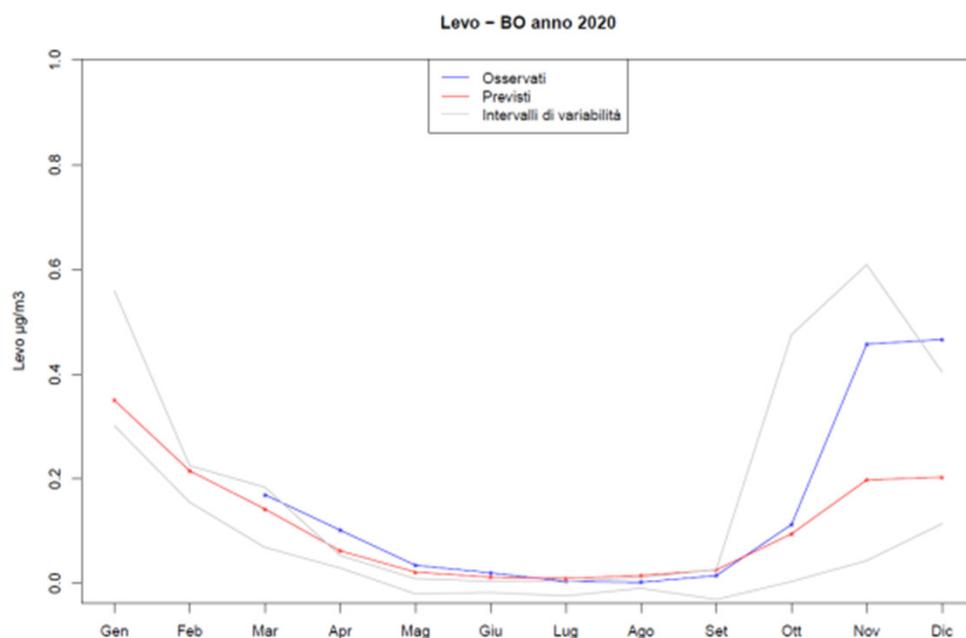
Il **carbonio elementare** presenta valori inferiori a quelli previsti per tutti i siti. La variazione delle concentrazioni osservate rispetto alle attese è compresa tra il -30% e il -65% circa nei siti durante il lockdown e tra il -20% e il -40% circa su tutto l'anno.

## Risultati: Ferro e Rame



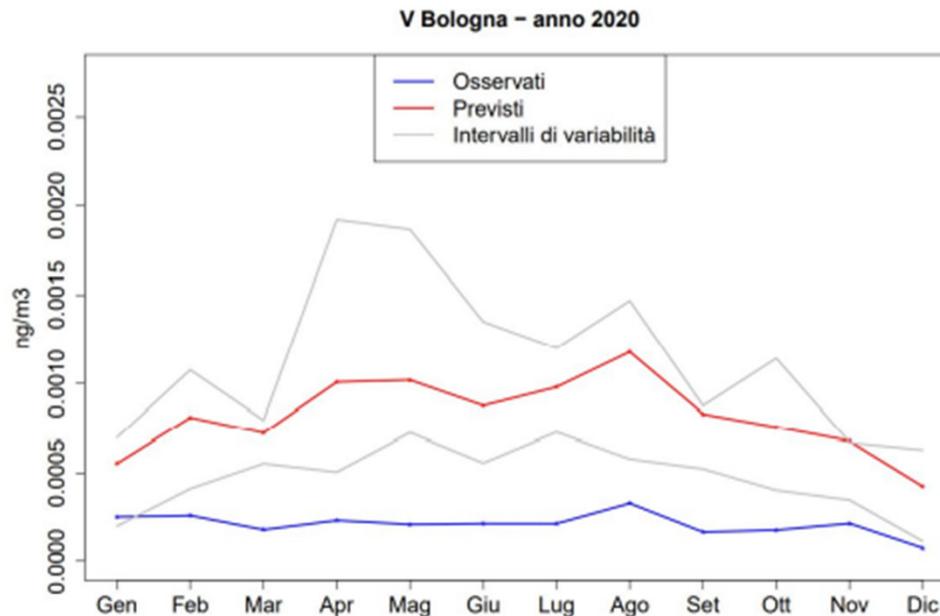
Ferro (Fe) e rame (Cu), che derivano in buona parte dal traffico, per un sito di traffico di Milano (PM10) mostrano un calo, rispettivamente, del 50% e del 70% nel periodo di lockdown e del 20% e del 30% circa per la media annua. Il calo durante il lockdown è visibile anche in un altro sito studiato che è un fondo urbano.

## Risultati: levoglucosano



Nel periodo di lockdown, il **levoglucosano**, marker della combustione della legna, mostra un aumento superiore al range di variabilità osservato negli anni precedenti. Per l'intera annualità non si evidenzia invece alcuna variazione maggiore della variabilità attesa.

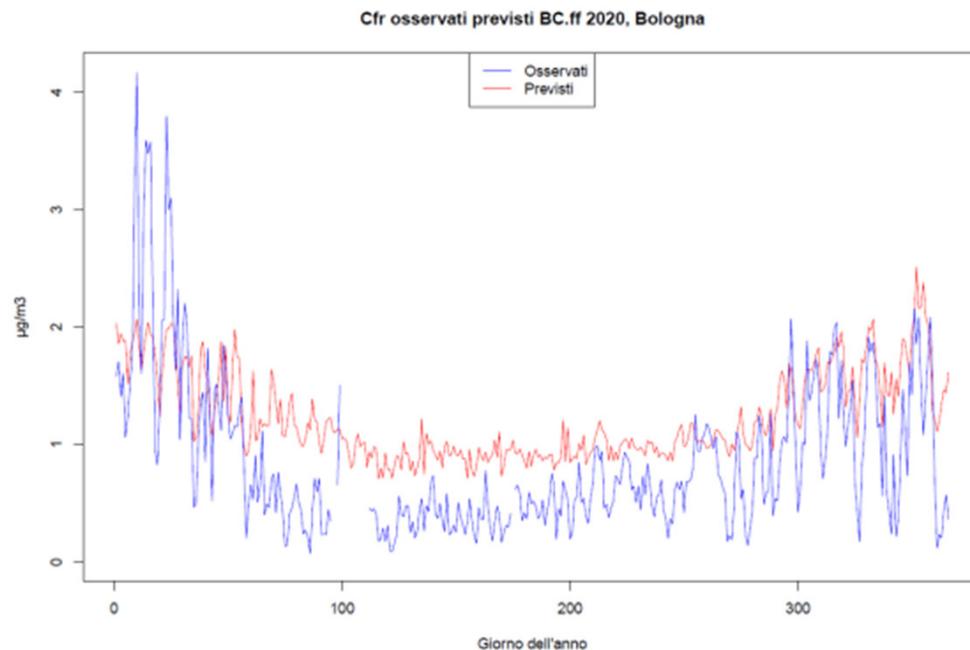
## Risultati: Vanadio



Il **vanadio** risulta un importante marker della combustione di oli pesanti. Le concentrazioni risultano molto più basse di quelle previste in tutto il 2020, anche prima del lockdown di marzo: i cali sono compresi tra il -70% e il -80%.

Si evidenzia anche una grande omogeneità tra i siti in cui viene misurato (PM2.5), probabilmente per il fatto che questo inquinante deriva prevalentemente da una sorgente di trasporto long range, la cui variazione impatta in modo simile su tutti i siti

## Risultati: black carbon

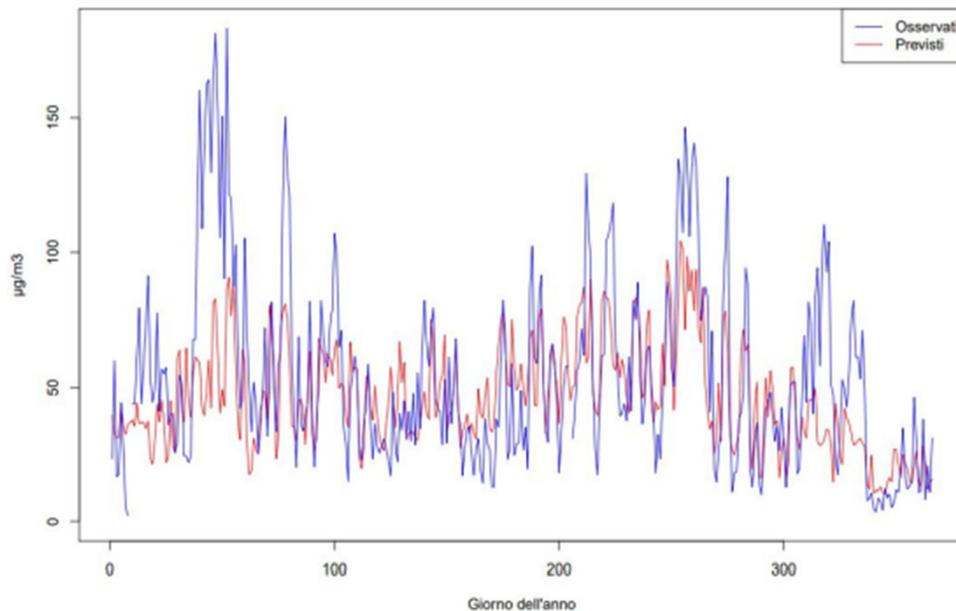


Si sono analizzate separatamente le due componenti del **black carbon** legate alla combustione di biomassa (BCbb) e al traffico (BCff). BCff evidenzia un netto calo in tutti i siti analizzati, sia nel periodo del lockdown che nei mesi estivi successivi; nei mesi autunnali sembra tornare in linea con i valori attesi per il periodo.

Il calo, nel periodo di lockdown è di circa il 60% nei siti di Bologna, Trento e Torino ed è pari al 50% circa nella stazione di Domodossola, che è un sito suburbano residenziale.

## Risultati: ammoniaca

Cfr osservati previsti NH3 2020, Corte de Cortesi



Per quel che riguarda l'**ammoniaca**, le variazioni tra dati previsti e osservati risultano paragonabili a quelle degli anni precedenti, sia durante il periodo del lockdown sia per l'intero anno 2020.

## Risultati

Le variazioni di **nitrato**, **ammonio** e **OC** ottenute per il 2020 (sia come intero anno che per il solo lockdown) sono invece comparabili con quelle calcolate per gli anni precedenti, di training, e non si possono pertanto attribuire alle azioni intraprese per combattere la diffusione del Covid-19.

## Conclusioni 1: risultati

Il confronto tra dato previsto in base alle condizioni meteorologiche in assenza di pandemia e dato osservato è affetto da una **importante incertezza** del metodo stesso, che in molti casi non permette di individuare la presenza o l'assenza di un effetto legato alle azioni intraprese per combattere la diffusione del Covid19.

Quando, però, questo effetto è individuabile, indica una **variazione sostanziale** della concentrazione degli inquinanti. Questo capita per tutti gli inquinanti che risultano traccianti del **traffico**: EC, BCff, Cu, Fe e, nei siti di traffico, persino NH<sub>3</sub>.

Per molti di questi inquinanti **il calo rispetto al valore BAU si osserva non soltanto durante il periodo di lockdown, ma anche per il resto dell'anno**, indicando che anche misure meno restrittive hanno portato ad un calo nelle concentrazioni degli inquinanti emessi principalmente dal traffico.

## Conclusioni 2: possibili impieghi della RF nello studio della qualità dell'aria

### TENERE IN CONSIDERAZIONE L'EFFETTO METEOROLOGICO

Riuscire a prevedere la concentrazione di un inquinante sulla base delle variabili meteorologiche permette di indagare:

- la presenza di una componente di lungo o medio periodo;
- l'effetto di altre misure (incentivi per cambio parco veicolare, limitazioni del traffico, apertura o chiusura di un impianto industriale, ...)

## Conclusioni 2: possibili impieghi della RF nello studio della qualità dell'aria

N.B.: Se non deve fare una previsione su anni fuori dalla **serie storica di addestramento** (il periodo non è perturbato e può essere utilizzato tutto per addestrare il modello) → non si verifica il problema del **trend** (bisogna se mai fare attenzione ad un **overfitting**)

## Conclusioni 2: possibili impieghi della RF nello studio della qualità dell'aria

UTILIZZARE LA RF COME METODO PER LA STIMA DI UNA SPECIE CHIMICA NON PER ELIMINARE L'EFFETTO METEOROLOGICO, MA AD ESEMPIO PER FARE UNA PREVISIONE, PER RICOSTRUIRE UN DATO MANCANTE, ...

N.B.: In questo caso si possono utilizzare anche

- **altre variabili esplicative**: componente autoregressiva, dati di altre centraline, dati di altri inquinanti,...
- ma anche **altre variabili dipendenti**: concentrazioni di inquinanti in altre matrici, concentrazioni polliniche,...



Fine

Grazie per l'attenzione